COMPOSITIONS THAT INCLUDE HEMAGGLUTININ, METHODS OF MAKING AND METHODS OF USE THEREOF

RELATED APPLICATIONS

This application is a continuation of and claims the benefit of U.S. Provisional Application Nos. 60/779,854, filed on March 7, 2006; 60/784,497, filed on March 20, 2006; 60/790,457, filed on April 7, 2006; 60/814,292, filed on June 16, 2006; 60/830,881, filed on July 14, 2006; 60/838,007, filed on August 16, 2006; and 60/856,451, filed on November 3, 2006. The entire teachings of the above applications are incorporated herein by reference.

BACKGROUND OF THE INVENTION

Viral influenza infection can lead to disease. Strategies to prevent and manage disease associated with viral influenza infection can include vaccines with inactivated viruses and drugs. However, such strategies can be costly to maintain supply with demand and, thus, be limited in supply; may result in variable protection and less than satisfactory alleviation of symptoms, thereby ineffectively preventing or treating illness and, in some instances death, consequent to disease associated with viral influenza infection. Thus, there is a need to develop new, improved and effective methods of treatment for preventing and managing disease associated with viral influenza infection.

SUMMARY OF THE INVENTION

The present invention relates to compositions, such as compositions that stimulate a protective immune response, and methods of making proteins that stimulate a protective immune response in a subject.

In one embodiment, the invention is a method of making a protein that stimulates a protective immune response in a subject, comprising the steps of separating a portion of a protein from a naturally occurring viral hemagglutinin to thereby form a protein portion, wherein the protein portion includes at least a portion of a globular head, and at least a portion of at least one secondary structure that causes the globular head to essentially retain its tertiary structure, and wherein the

protein portion lacks a membrane fusion domain, a transmembrane domain and a cytoplasmic domain; transforming a nucleic acid sequence encoding the protein portion into a prokaryotic host cell; and culturing the prokaryotic host cell to thereby make the protein that stimulates a protective immune response in a subject.

In another embodiment, the invention is a method of making a protein that stimulates a protective immune response in a subject, comprising the steps of separating a portion of a protein from a naturally occurring viral hemagglutinin to thereby form a protein portion, wherein the protein portion includes at least a portion of a globular head, and at least a portion of at least one secondary structure that causes the globular head to essentially retain its tertiary structure, and wherein the protein portion lacks a membrane fusion domain, a transmembrane domain and a cytoplasmic domain; transfecting a nucleic acid sequence encoding the protein portion into a eukaryotic host cell, wherein the eukaryotic host cell is not a *Pichia pastoris* eukaryotic host cell; and culturing the eukaryotic host cell to thereby make the protein that stimulates a protective immune response in a subject.

In another embodiment, the invention is a method of making a protein that stimulates a protective immune response in a subject, comprising the steps of separating a portion of a protein from a naturally occurring viral hemagglutinin to thereby form a protein portion, wherein the protein portion includes at least a portion of a globular head, and at least a portion of at least one secondary structure that causes the globular head to essentially retain its tertiary structure, and wherein the protein portion lacks a membrane fusion domain, a transmembrane domain and a cytoplasmic domain; transfecting a nucleic acid sequence encoding the protein portion into a eukaryotic host cell, wherein the eukaryotic host cell is not a *Drosophila melanogaster* eukaryotic host cell; and culturing the eukaryotic host cell to thereby make the protein that stimulates a protective immune response in a subject.

In a further embodiment, the invention is a method of making a protein that stimulates a protective immune response in a subject, comprising the steps of separating a portion of a protein from a naturally occurring viral hemagglutinin to thereby form a protein portion, wherein the protein portion includes at least a portion of a globular head, and at least a portion of at least one secondary structure that

causes the globular head to essentially retain its tertiary structure, and wherein the protein portion lacks a membrane fusion domain, a transmembrane domain and a cytoplasmic domain; transfecting a nucleic acid sequence encoding the protein portion into a eukaryotic host cell, wherein the eukaryotic host cell is not an insect eukaryotic host cell; and culturing the eukaryotic host cell to thereby make the protein that stimulates a protective immune response in a subject.

In a further embodiment, the invention is a method of making a protein that stimulates a protective immune response in a subject, comprising the steps of separating a portion of a protein from a naturally occurring viral hemagglutinin to thereby form a protein portion, wherein the protein portion includes at least a portion of a globular head, and at least a portion of at least one secondary structure that causes the globular head to essentially retain its tertiary structure, and wherein the protein portion lacks a membrane fusion domain, a transmembrane domain and a cytoplasmic domain; transfecting a nucleic acid sequence encoding the protein portion into a eukaryotic host cell, wherein the eukaryotic host cell is not a stably transformed insect host cell; and culturing the eukaryotic host cell to thereby make the protein that stimulates a protective immune response in a subject.

In still another embodiment, the invention is a method of making a protein that stimulates a protective immune response in a subject, comprising the steps of separating a portion of a protein from a naturally occurring viral hemagglutinin to thereby form a protein portion, wherein the protein portion includes at least a portion of a globular head, and at least a portion of at least one secondary structure that causes the globular head to essentially retain its tertiary structure, and wherein the protein portion lacks a membrane fusion domain, a transmembrane domain and a cytoplasmic domain; transfecting a nucleic acid sequence encoding the protein portion into a eukaryotic host cell, wherein the eukaryotic host cell is neither a *Picia pastoris* eukaryotic host cell nor a stably transfected insect host cell; and culturing the eukaryotic host cell to thereby make the protein that stimulates a protective immune response in a subject.

In yet another embodiment, the invention is a method of stimulating protective immunity in a subject, comprising the step of administering to the subject a composition that includes a protein made by a method comprising the steps of

separating a portion of a protein from a naturally occurring viral hemagglutinin to thereby form a protein portion, wherein the protein portion includes at least a portion of a globular head, and at least a portion of at least one secondary structure that causes the globular head to essentially retain its tertiary structure, and wherein the protein portion lacks a membrane fusion domain, a transmembrane domain and a cytoplasmic domain; transforming a nucleic acid sequence encoding the portion into a prokaryotic host cell; and culturing the prokaryotic host cell to thereby make the protein that stimulates a protective immune response in a subject.

In an additional embodiment, the invention is a method of stimulating protective immunity in a subject, comprising the step of administering to the subject a composition that includes a protein portion of a naturally occurring viral hemagglutinin, wherein the protein portion includes at least a portion of a globular head, and at least a portion of one secondary structure that causes the globular head to essentially retain its tertiary structure and wherein the protein portion lacks a membrane fusion domain, a transmembrane domain and a cytoplasmic domain.

In still another embodiment, the invention is a method of making a viral hemagglutinin protein that stimulates a protective immune response in a subject, comprising the steps of separating a portion of a protein from a naturally occurring viral hemagglutinin to thereby form a protein portion, wherein the protein portion includes at least a portion of a globular head and at least a portion of at least one secondary structure that causes the globular head to essentially retain its tertiary structure, and wherein the protein portion lacks a membrane fusion domain, a transmembrane domain and a cytoplasmic domain; transfecting a nucleic acid sequence encoding the portion in a eukaryotic host cell, wherein the eukaryotic host cell is neither a *Pichia pastoris* eukaryotic host cell nor a stably transfected insect host cell; and culturing the eukaryotic host cell to thereby make the protein that stimulates a protective immune response in a subject.

In another embodiment, the invention is a method of making a protein that stimulates a protective immune response in a subject, comprising the steps of separating a portion of a protein from a naturally occurring viral hemagglutinin to thereby make a protein portion, wherein the protein portion includes at least a portion of a globular head, and at least a portion of at least one secondary structure

that causes the globular head to essentially retain its tertiary structure, and wherein the protein portion lacks a membrane fusion domain, a transmembrane domain and a cytoplasmic domain; infecting a nucleic acid sequence encoding the protein portion into an insect cell host cell; and culturing the insect host cell to thereby make the protein that stimulates a protective immune response in a subject.

In another embodiment, the invention is a method of making a viral hemagglutinin protein that stimulates a protective immune response in a subject, comprising the steps of transforming a prokaryotic host cell with a nucleic acid sequence encoding at least one viral hemagglutinin that lacks a transmembrane domain and a cytoplasmic domain; and culturing the prokaryotic cell to thereby make the protein.

In a further embodiment, the invention is a method of stimulating protective immunity in a subject, comprising the step of administering to the subject a composition that includes a protein made by a method comprising the steps of transforming a prokaryotic host cell with a nucleic acid sequence encoding at least one viral hemagglutinin that lacks a transmembrane domain and a cytoplasmic domain; and culturing the prokaryotic host cell to thereby make the protein.

In another embodiment, the invention is a method of stimulating protective immunity in a subject, comprising the step of administering to the subject a composition that includes a protein having at least one viral hemagglutinin that lacks a transmembrane domain and a cytoplasmic domain, wherein the protein was expressed in a prokaryotic cell.

In an additional embodiment, the invention is a composition comprising at least a portion of at least one pathogen-associated molecular pattern and a portion of a protein of a naturally occurring viral hemagglutinin, wherein the portion of the naturally occurring viral hemagglutinin includes at least a portion of a globular head and at least a portion of at least one secondary structure that causes the globular head to essentially retain its tertiary structure, and wherein the portion of the naturally occurring viral hemagglutinin lacks a membrane fusion domain, a transmembrane domain, and a cytoplasmic domain.

Another embodiment of the invention is a composition comprising a flagellin component that is at least a portion of a flagellin, wherein the flagellin component

includes at least one cysteine residue and whereby the flagellin component activates a Toll-like Receptor 5.

In still another embodiment, the invention is a composition comprising a Toll-like Receptor agonist component that is at least a portion of a Toll-like Receptor agonist, wherein the Toll-like Receptor agonist component includes at least one cysteine residue in a position where a cysteine residue does not occur in the native Toll-like Receptor agonist, whereby the Toll-like Receptor agonist component activates a Toll-like Receptor.

In an additional embodiment, the invention is a composition comprising a flagellin component that is at least a portion of a flagellin, wherein at least one lysine of the flagellin component has been substituted with at least one arginine, whereby the flagellin component activates a Toll-like Receptor 5.

In yet another embodiment, the invention is a composition comprising a flagellin component that is at least a portion of a flagellin, wherein at least one lysine of the flagellin component has been substituted with at least one serine residue, whereby the flagellin component activates a Toll-like Receptor 5.

Another embodiment of the invention is a composition comprising a flagellin component that is at least a portion of a flagellin, wherein at least one lysine of the flagellin component has been substituted with at least one histidine residue, whereby the flagellin component activates a Toll-like Receptor 5.

A further embodiment of the invention is a method of stimulating an immune response in a subject, comprising the step of administering to the subject a composition that includes a flagellin component that is at least a portion of a flagellin, wherein the flagellin component includes at least one cysteine residue and whereby the flagellin component activates a Toll-like Receptor 5.

In still another embodiment, the invention is a method of stimulating an immune response in a subject, comprising the step of administering to the subject a composition that includes a flagellin component that is at least a portion of a flagellin, wherein at least one lysine of the flagellin component has been substituted with at least one arginine, whereby the flagellin component activates a Toll-like Receptor 5.

An additional embodiment of the invention is a method of stimulating an immune response in a subject, comprising the step of administering to the subject a composition that includes a flagellin component that is at least a portion of a flagellin, wherein at least one lysine of the flagellin component has been substituted with at least one serine residue, whereby the flagellin component activates a Toll-like Receptor 5.

In another embodiment, the invention is a method of stimulating an immune response in a subject, comprising the step of administering to the subject a composition that includes a flagellin component that is at least a portion of a flagellin, wherein at least one lysine of the flagellin component has been substituted with at least one histidine residue, whereby the flagellin component activates a Toll-like Receptor 5.

In yet another embodiment, the invention is a method of stimulating an immune response in a subject, comprising the step of administering to the subject a composition that includes a Toll-like Receptor agonist component that is at least a portion of a Toll-like Receptor agonist, wherein the Toll-like Receptor agonist component includes at least one cysteine residue in a position where a cysteine residue does not occur in the native Toll-like Receptor agonist, whereby the Toll-like Receptor agonist component activates a Toll-like Receptor agonist.

The methods and composition of the invention can be employed to stimulate an immune response, in particular, a protective immune response, in a subject. Advantages of the claimed invention include, for example, cost effective methods and compositions that can be produced in relatively large quantities for use in the prevention and treatment of disease associated with viral influenza infection. The claimed methods and compositions can be employed to prevent or treat viral influenza infection and, therefore, avoid serious illness and death consequent to viral influenza infection.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 depicts a ribbon diagram of PR8 influenza A hemagglutin (HA) crystal structure (1RU7). Schematically presentations from left to right are HA trimer, monomer, the globular head domain and the three choices of vaccine

candidates. Molecules or domains of interest are highlighted by dotted circle, such as the monomer in trimer structure and the globular head domain within the monomer structure. The domain boundaries of three vaccine candidates are marked by the cross signs. The residue numbers of each selection are also labeled in the individual vaccine presentation.

Figure 2 depicts a hydrophobicity plot analysis of the B/Lee/40 HA (SEQ ID NO: 36) using ProtScale (http://ca.expasy.org/tools → Primary structure analysis → ProtScale) to confirm that the selected boundaries were in the hydrophilic regions of the protein.

Figure 3 depicts a sandwich ELISA analysis of STF2.HA B strain proteins. ELISA plates were coated with antibody to flagellin, and the indicated proteins were incubated to allow capture. Proteins were detected using ferret antisera against B/Malaysia/2506/2004 and enzyme-labeled goat anti-ferret antibodies.

Figure 4 depicts the TLR bioactivity of STF2.HA B strain proteins. HEK293 (TLR5+) cells were incubated with the indicated proteins overnight, and the cell culture supernatants were harvested and assessed for IL-8 by ELISA.

Figure 5 depicts the TLR bioactivity of STF2.HA B strain proteins. HEK293 (TLR5+) cells were incubated with the indicated proteins overnight, and the cell culture supernatants were harvested and assessed for IL-8 by ELISA.

Figure 6 depicts the dose-dependent antibody response to STF2.HA1-2(PR8) (SEQ ID NO: 90) in BALB/c mice. Mice (10/group) were immunized as indicated on days 0 and 14, and bled on days 12 and 21. Anti-HA IgG responses on Day 28 were examined by ELISA. Convalescent antisera (x) were included as a positive control. The data depict the mean±SD of 10 individual sera per group.

Figure 7 depicts serum reactivity with influenza infected cells *in vitro*. The day 21 serum samples depicted in Figure 6 were incubated with mock- and PR/8/34-infected MDCK cells. The data represent the OD_{450} mean ± SD of 10 individual sera/group.

Figure 8A depicts survival of BALB/c mice immunized with STF2.HA1-2(PR8) (SEQ ID NO: 90). Mice from Figures 8 and 9 were challenged on day 28 with an LD_{90} ($8x10^3$ EID) of influenza A PR/8/34 administered intranasally (Day 0 post-challenge).

Figure 8B depicts weights of BALB/c mice immunized with STF2.HA1-2(PR8) (SEQ ID NO: 90). Mice from Figures 6 and 7 were challenged on day 28 with an LD_{90} ($8x10^3$ EID) of influenza A PR/8/34 administered intranasally (Day 0 post-challenge). Graph reflects the average weight per group based on individual animals measured daily for 19 days.

Figure 8C depicts clinical Score of BALB/c mice immunized with STF2.HA1-2(PR8) (SEQ ID: NO 90). Mice from Figures 6 and 7 were challenged on day 28 with an LD_{90} ($8x10^3$ EID) of influenza A PR/8/34 administered intranasally (Day 0 post-challenge). Graph reflects the average clinical score per group based on individual animals measured daily for 19 days. Clinical scores were assessed as follows: 4 pts = healthy, 3 pts = reduced grooming, 2 pts = reduced physical activity and 1 pt = moribund.

Figure 9 depicts neutralization activity of immune sera from BALB/c mice immunized with the indicated doses of STF2.HA1-2(PR8) (SEQ ID NO: 90) on days 0 and 14, and bled on day 21. The endpoint titer of each sample was calculated as the titer that inhibited virus-mediated lysis of the cells by at least 50%. Geometric mean titers ± SD were calculated for each dose group and plotted.

Figures 10A and 10B depict anti-HA antibody response of BALB/c mice immunized with 3 μg of STF2.HA1-2(PR8) (SEQ ID NO: 90), or 3 or 0.3 μg of STF2.HA1-2(VN) (SEQ ID NO: 95) s.c. on days 0 and 14. On day 21, sera were isolated and examined for reactivity with (Figure 10A) HA purified from Influenza A/Viet Nam/1203/2004 (obtained from BEI Resources Cat# NR-660)) and (Figure 10B) recombinant flagellin (STF2) (SEQ ID NO: 96) by ELISA.

Figure 11 depicts a flow chart of the steps required to express pFastBac constructs using Bac-to-Bac *Baculovirus* Expression system.

Figures 12A and 12B depict an anti-HA specific IgG responses in BALB/c mice immunized with STF2.HA1-1 fusion proteins. Mice (10/group) were immunized with indicated STF2.HA1-1 fusion proteins on days 0 and 14, and bled on day 21. Anti-HA and anti-flagellin IgG responses were examined by ELISA. Convalescent antisera were included as a positive control. The data depict the mean±SD of 10 individual sera per group.

Figure 13A depicts survival of BALB/c mice immunized with recombinant STF2.HA1-1 proteins. On day 28, animals in Figure 12 were challenged i.n. with an LD_{90} ($8x10^3$ EID) of influenza A PR/8/34. The survival of individual mice was monitored for 21 days post challenge.

Figure 13B depicts weight of BALB/c mice immunized with recombinant STF2.HA1-1 proteins. On day 28, animals in Figure 12 were challenged i.n. with an LD_{90} ($8x10^3$ EID) of influenza A PR/8/34. Graph reflects the average weight per group based on individual animals measured daily for 21 days.

Figure 13C depicts clinical assessment of BALB/c mice immunized with recombinant STF2.HA1-1 proteins. On day 28, animals in Figure 12 were challenged intranasally (i.n.) with an LD_{90} ($8x10^3$ EID) of influenza A PR/8/34. Graph reflects the average weight per group based on individual animals measured daily for 21 days. Clinical scores were assessed as follows: 4 pts = healthy, 3 pts = reduced grooming, 2 pts = reduced physical activity and 1 pt = moribund.

Figure 14 depicts anti-HA antibody titers of mice following immunization with 10 μg STF2.HA1-2(VN).

Figure 15 depicts an anti-HA antibody titers of mice following immunization with 3 μg STF2.HA1-2(VN).

Figure 16 depicts an anti-HA antibody titers of mice following immunization with 1 μg STF2.HA1-2(VN).

Figure 17 depicts a survival of mice following challenge with 10 LD_{90} of influenza A/Viet Nam/1203/04

Figure 18 depicts the amino acid sequence of full-length flagellin (STF2) highlighting important domains and residues. Gray = experimentally-defined TLR5 binding site (Smith, et al., 2003); K= lysine residues; underline = flagellin sequences corresponding to STF2Δ construct (linker not shown).

Figure 19 depicts a space-filling model of flagellin (STF2) with the TLR5 activation site (left panel) and lysine residues (right panel) highlighted in gray.

Figure 20 depicts a space-filling model of flagellin (STF2) with residues retained in STF2Δ highlighted in black.

Figure 21 depicts a schematic depiction of the two-step PCR cloning of pMT/STF2Δ.

Figure 22 depicts recognition of flagellin epitopes of STFΔ proteins in ELISA.

Figure 23 depicts TLR5-dependent secretion of IL-8 following stimulation of HEK293 cells with the indicated STF2Δ proteins.

Figure 24 depicts the S200 chromatogram of STF2Δ.HingeCys:CysH1C1. Void = column void volume; conjugate = elution peak of protein: peptide conjugate; A260, A280 and conductivity are indicated.

Figure 25 depicts TLR5-dependent secretion of IL-8 following stimulation of RAWhTLR5 cells (closed symbols) or TLR5-negative RAW cells (open symbols) with STF2Δ.HingeCys (squares), STF2Δ.HingeCys:CysH1C1 conjugate (circles), or STF2.OVA (triangles).

Figure 26 depicts the S200 chromatogram of STF2.4xH1C1. Void = column void volume; conjugate = elution peak of protein: peptide conjugate; A260, A280 and conductivity are indicated.

Figure 27 depicts an anti-H1C1 antibody response of BALB/c mice immunized with native H1C1 peptide or Pam3Cys.H1C1 peptide.

Figure 28 depicts a survival plot of BALB/c mice immunized with native H1C1 peptide or Pam3Cys.H1C1 peptide then challenged with an LD_{90} of influenza A/Puerto Rico/8/34 virus.

Figure 29 depicts the amino acid sequence of *Salmonella typhimurium* flagellin type 2 (fljB/STF2) with the hinge region underlined (SEQ ID NO: 498).

Figure 30 depicts the nucleic acid sequence (SEQ ID NO: 499) encoding SEQ ID NO: 498. The nucleic acid sequence encoding the hinge region is underlined.

Figure 31 depicts the amino acid sequence of fljB/STF2 without the hinge region (also referred to herein as "fljB/STF2Δ" or "STF2Δ") (SEQ ID NO: 500).

Figure 32 depicts the nucleic acid sequence (SEQ ID NO: 501) encoding SEQ ID NO: 500.

Figure 33 depicts the amino acid sequence of *E. coli* flagellin fliC (also referred to herein as "*E. coli* fliC") with the hinge region underlined (SEQ ID NO: 502).

Figure 34 depicts the nucleic acid sequence (SEQ ID NO: 503) encoding SEQ ID NO: 502. The nucleic acid sequence encoding the hinge region is underlined.

Figure 35 depicts the amino acid sequence of *S. muenchen* flagellin fliC (also referred to herein as "*S. muenchen* fliC") with the hinge region underlined (SEQ ID NO: 504).

Figure 36 depicts the nucleic acid sequence (SEQ ID NO: 505) encoding SEQ ID NO:504. The nucleic acid sequence encoding the hinge region is underlined.

Figure 37 depicts the amino acid sequence of pMT/STF2. The linker is underlined and the sequence of the BiP secretion signal is bolded (SEQ ID NO: 506).

Figure 38 depicts the nucleic acid sequence (SEQ ID NO: 507) of SEQ ID NO: 506. The nucleic acid sequence encoding the linker is underlined and the nucleic acid sequence encoding the BiP sequence is bolded.

Figure 39 depicts a Pam3Cys.M2e fusion protein. The amino acid sequence (SEQ ID NO: 510) of M2e is shown in bold type.

Figure 40 depicts the activation of an antigen-presenting cell (APC) by Toll-like Receptor (TLR) signaling.

Figures 41A and 41B depict plasmid constructs to express an amino-terminus of an M2 (e.g., SEQ ID NOS: 510, 554) of H1 and H5 (SEQ ID NO: 536) influenza A viral isolates. pMT: metallothionein promoter-based expression vector. BiP: secretion signal sequence of immunoglobulin-binding protein. STF2: full-length flagellin of *S. typhimurium*. STF2Δ: hinge region-deleted STF2. MCS: multiple cloning site.

Figure 42 depicts plasmid constructs designed to express HA of H1 and H5 influenza A virus isolates. AOX1: AOX1 promoter of pPICZα expression vector (Invitrogen Corporation, Carlsbad, CA). αf: secretion signal sequence of yeast. STF2: full-length flagellin of *S. typhimurium*. STF2Δ: hinge region-deleted STF2. MCS: multiple cloning site.

Figure 43 depicts the amino acid sequence (SEQ ID NO: 561) of HA (PR8).

Figure 44 depicts the nucleic acid sequence (SEQ ID NO: 562) encoding SEQ ID NO: 561.

Figure 45 depicts the amino acid sequence (SEQ ID NO: 563) of *E. coli* fliC without the hinge region.

Figure 46 depicts the amino acid sequence of pMT/STF2Δ (SEQ ID NO: 585). The linker sequence is underlined and the BiP secretion signal is bolded.

Figure 47 depicts the nucleic acid sequence (SEQ ID NO: 586) encoding SEQ ID NO: 585. The nucleic acid sequence encoding the linker is underlined and the nucleic acid sequence encoding the BiP secretion signal is bolded.

Figure 48 depicts the amino acid sequence (SEQ ID NO: 595) of the *Salmonella muenchen* fliC without the hinge region, which is also referred to herein as "*S. muenchen* fliCΔ."

Figure 49 depicts the nucleic acid sequence of *Salmonella muenchen* fliC (SEQ ID NO: 596) encoding SEQ ID NO: 595.

Figure 50 depicts IL-8 secretion following stimulation of TLR5+ cells.

Figure 51 depicts TNF secretion following stimulation of TLR2+ cells.

Figure 52 depicts M2e-specific IgG.

Figure 53 depicts the OVA-specific IgG.

Figure 54 depicts the M2e-specific IgG serum titers.

Figure 55 depicts the M2e-specific serum IgG titer post-boost.

Figure 56 depicts the Pam3Cys.M2e dose response.

Figure 57 depicts the M2e-specific serum IgG titer.

Figure 58 depicts the rabbit IgG response to M2e.

Figure 59 depicts the immunogenicity of STF2.4xM2e in a rabbit 14 days post-prime.

Figure 60 depicts the survival of BALB1c mice following challenge. Mice were immunized on days 0 and 14 as indicated in the legend and challenged on day 28 by intranasal administration of an LD_{90} of influenza A/Puerto Rico/8/34 virus (Day 0 post-challenge). Mice were monitored for survival for 21 days post-challenge.

Figure 61 depicts fusion contructs in a pET24 vector. T7:T7 promoter; *lac*O: lac operator; STF2: *Salmonella typhimurium* flagellin; STF2Δ= STF2 with the hinge

region deleted; $EIII^{+}$ is domain III of a West Nile envelope protein with 6 amino acids of domain I amino acid.

Figures 62A and 62B depict TLR5 bioactivity of STF2.EIII+ (SEQ ID NOS: 657, 658) and STF2ΔEIII+ (SEQ ID NOS: 673, 674) fusion proteins. Serial dilutions of purified proteins were added to HEK293 (TLR5+) cells overnight and IL-8 content of the supernatants measured by ELISA. Purified STF2.OVA was used as a positive control (Figure 129A). The TLR2 agonist Pam3CSK4 was used as a negative control (Figure 129B).

Figure 63 depicts STF2Δ.EIII+ antigenic epitopes assessed by ELISA. Plates were coated with full-length WNE (open bars) (SEQ ID NO: 642) or STF2Δ.EIII+ (SEQ ID NOS: 673,674) and probed with the indicated antibodies (mAb). Poly = polyclonal antiserum to WNE; 3D9 through 7H2 = neutralizing monoclonal antibodies to WNE epitopes; anti-flagellin = monoclonal antibody to flagellin.

Figures 64A, 64B, 64C and 64D depict reactivity of STF2.E (SEQ ID NOS: 761, 762); STF2.EIII+ (SEQ ID NOS:657, 658) and STF2Δ.EIII+ (SEQ ID NOS:673,674) fusion proteins with antibodies to WNE and flagellin. Plates were coated with fusion proteins, blocked and incubated with antibodies to WNE or flagellin. Antibody reactivity was detected following incubation with HRP-labeled species specific IgG. Plates were developed in the presence of TMB substrate and O.D.450/650 using a TECAN plate reader and Magellian software.

Figure 65 depicts IgG serum following injection with fusion proteins. Mice were immunized with either PBS, *Drosophila* conditioned medium containing STF2.E (CM, positive control), 25 μg of STF2Δ.EIII+ (SEQ ID NOS:673, 674) i.p., 25 μg STF2Δ.EIII+ s.c., 25 μg STF2.EIII+ (SEQ ID NO:659,658) i.p., 25 μg STF2.EIII+ (SEQ ID NOS:657,658) or 25μg STF2.E (SEQ ID NOS: 761, 762). On day 35, immunized animals were challenged with WNV. Sera from individual mice (day 35) were characterized by direct ELISA to determine IgG levels. Purified WNV-E protein (SEQ ID NO:642) was used as the antigen in this assay. This antigen (60) was produced in *Drosophila* as a his-tagged protein.

Figure 66 depicts STF2Δ.EIII+ (SEQ ID NOS: 673 674) and STF2.EIII+ (SEQ ID NOS: 657, 658) protective immunity to WNV viral challenge. Mice were

immunized and challenged with a lethal dose of WNV strain 2741 on day 35. Survival was monitored for 21 days.

Figure 67 depicts IgG sera titers following immunization with fusion proteins. STF2Δ.EIII+ proteins induce WNV-specific IgG antibodies. Mice were immunized s.c. on days 0, 14 and 28 with PBS alone or about 25 μg of STF2Δ.EIII+ (SEQ ID NOS: 673, 674) (045 [positive control]), STF2Δ.EIII+ (067, trimer), STF2Δ.EIII+ (070, monomer) or STF2Δ.EIIIs+ (SEQ ID NOS: 675, 676) (069). On day 35 sera from individual mice were characterized by direct ELISA to determine IgG levels. Purified WNV-E protein (060, produced in *Drosophila* as a his-tagged protein) was used as the antigen in this assay.

Figure 68 depicts STF2Δ.EIII+ (SEQ ID NOS: 673, 674) and STF2Δ.EIIIs+ (SEQ ID NOS: 675, 673) protective immunity in mice from WNV lethal challenge. On day 38 following immunization with fusion proteins, all groups were challenged with a lethal dose of WNV strain 2741 and survival was monitored for 21 days. Survival for each group (10 mice/group) is indicated as a percentage.

Figure 69 depicts competition assays. Serial dilutions (five fold starting at 1:25) of antisera from immunized animals were incubated with biotinylated WNE protein (SEQ ID NO: 642) and then added to the wells of ELISA plates coated with mAb 7H2 at about 2mg/ml. Wells were developed using avidin-HRP to determine inhibition of West Nile protein binding as a results of competition with mAb 7H2.

Figure 70 depicts epitope mapping of the antibody response induced by STF2Δ.EIII+ (SEQ ID NOS: 675,676) fusion proteins. Immune sera from animals immunized with indicated STF2Δ-fusion proteins (E2-21, E27-E52, Figure 142) were examined for the ability to recognize overlapping peptides corresponding to the junction of domains I and III of the WNV envelope protein.

Figure 71 depicts epitope mapping of the antibody response induced by STFΔ.EIIIs+ (SEQ ID NOS: 675, 676) E-21 (envelope protein) epitope fusion proteins. Immune sera from animals immunized with the indicated STF2Δ-fusion proteins (E2-21, E2-21-1(S,C), E2-21-2(C,S), E2-21-2(C,S) and E2-21-4 through E2-21-24, see Figure 139) were evaluated to identify the residues defining the E-21 epitope of West Nile envelope protein. Data reflects the response of sera to E-21 following the substitution of cysteine with serine (indicated by C,S); and the

sequential replacement of amino acids with alanine. The peptides tested are listed in Figure 139.

Figure 72 depicts Pam3Cys.WNV001 (SEQ ID NO: 771) inducing EIII specific IgG antibodies. Mice were immunized s.c. on days 0, 14 and 28 with PBS alone, 22mg of unmodified WNV001 (SEQ ID NO: 771) or 30 μg of Pam3Cys.WNV001. On day 35 sera from individual mice were characterized by direct ELISA to determine IgG levels to synthetic WNV001 peptide.

Figure 73 depicts the amino acid sequences (SEQ ID NOS: 691-698) of the EI/EIII junction for West Nile, Japanese encephalitis and Dengue (serotypes 1 through 4) viruses. The West Nile epitope identified using antisera from STF2Δ.EIIIs+ immunized animals is underlined. This sequence corresponds to peptide E2-21 (SEQ ID NO: 728).

Figure 74 depicts a tripalmitoylated peptide.

Figure 75 depicts the D1 domain, D2 domain, TLR5 activation domain and hypervariable (D3 domain) of flagellin.

Figure 76 depicts the D1 domain, D2 domain, TLR5 activation domain and hypervariable (D3 domain) of flagellin (Yonekura, et al. Nature 424, 643-650 (2003)).

Figure 77 depicts the D0, D1, D2 and D3 domains of flagellin and regions in the D2 and D3 domains suitable for insertion or substitution with antigens (e.g., maturational cleavage site, HA, M2e).

Figure 78 depicts the amino acid sequence of *Pseudomonas aeruginosa* flagellin (SEQ ID NO: 815). Lysine residues are indicated by the asterisks.

Figure 79 depicts the amino acid sequence of a *S. typhimurium* flagellin (SEQ ID NO: 816). Lysine residues are indicated by the asterisks.

Figure 80 depicts the amino acid sequence (SEQ ID NO: 820) of *Listeria monocytogenes* flagellin (GenBank Accession No: Q92DW3). Lysine residues are indicated by the asterisks.

DETAILED DESCRIPTION OF THE INVENTION

The features and other details of the invention, either as steps of the invention or as combinations of parts of the invention, will now be more particularly

described and pointed out in the claims. It will be understood that the particular embodiments of the invention are shown by way of illustration and not as limitations of the invention. The principle features of this invention can be employed in various embodiments without departing from the scope of the invention.

The invention generally is directed to methods of making compositions, and to compositions, that stimulate an immune response in a subject, such as a protective immune response, and methods of treatment, such as by administering the composition to a subject.

"Stimulating an immune response," as used herein, refers to the generation of antibodies and/or T-cells to at least a portion of an antigen, such as the protein portions of hemagglutinin (HA) (e.g., HA1-1, HA1-2 proteins) described herein. The antibodies and/or T-cells can be generated to at least a portion of an influenza viral protein (e.g., HA and M2 proteins of influenza A, B and/or C).

Stimulating an immune response in a subject can include the production of humoral and/or cellular immune responses that are reactive against the antigen, such as a viral protein, in particular, an influenza viral protein.

The compositions of the invention for use in methods to stimulate immune responses in subjects, can be evaluated for the ability to stimulate an immune response in a subject using well-established methods. Exemplary methods to determine whether the compositions of the invention stimulate an immune response in a subject, include measuring the production of antibodies specific to the antigen (e.g., IgG antibodies) by a suitable technique such as, ELISA assays; the potential to induce antibody-dependent enhancement (ADE) of a secondary infection; macrophage-like assays; neutralization assessed by using the Plaque Reduction Neutralization Test ($PRNT_{80}$); and the ability to generate serum antibodies in non-human models (e.g., mice, rabbits, monkeys) (Putnak, *et al., Vaccine 23*:4442-4452 (2005)).

"Stimulates a protective immune response," as used herein, means administration of the compositions of the invention, such as hemagglutinin (HA) proteins (e.g., HA1-1, HA1-2 proteins described herein), results in production of antibodies to the protein to thereby cause a subject to survive challenge by an otherwise lethal dose of a viral protein, such as viral HA. Techniques to determine a

lethal dose of a virus (e.g., an influenza virus) are known to one of skill in the art (see, for example, WHO/CDS/CSR/NCS2002.5 "WHO Manual on Animal Influenza Diagnosis and Surveillance" World Health Organization, Dept of Communicable Disease Surveillance and Response, WHO Global Influenza Programme; Harmon, M.W., *et al., J. Clin. Microbiol. 26*:333-337 (1988); Reed, L.J., *et al., Am. J. Hyg. 27*:493-497 (1938); Rose, T., *et al., J. Clin. Microbiol. 37*:937-943 (1999); Walls, H.H. *et al., J. Clin. Microbiol. 23*:240-245 (1986); Current Protocols in Immunology, 19.11.1-19.11.32, Cottey, R., *et al.,* John Wiley & Sons, Inc (2001)). Exemplary techniques for determining a lethal dose can include administration of varying doses of virus and a determination of the percent of subjects that survive following administration of the dose of virus (e.g., LD_{10}, LD_{20}, LD_{40}, LD_{50}, LD_{60}, LD_{70}, LD_{80}, LD_{90}). For example, a lethal dose of a virus that results in the death of 50% of a population of subjects is referred to as an "LD_{50}"; a lethal dose of a virus that results in the death of 80% of a population of subjects is referred to herein as "LD_{80}"; a lethal dose of a virus that results in death of 90% of a population of subjects is referred to herein as "LD_{90}."

For example, determination of the LD_{90} can be conducted in subjects (e.g., mice) by administering intranasally varying doses (*e.g.*, dilutions, such as log and half-log dilutions of 8×10^3 egg-infectious doses (EID)) followed by an assessment of the survival of the subjects about 14 days to about 21 days after infection with the virus. Protective immunity can be assessed by physical appearance of the subject, general demeanor (active), weight (initial loss of weight followed by return to a weight about the weight of the subject prior to infection with the virus) and survival after about 14 to about 21 days following infection with the virus.

Assessment of stimulation of protective immunity can also be made by employing assays that assess the ability of the antibodies produced in response to the compositions of the invention (e.g., a portion of the protein of the naturally occurring virus, such as a protein portion of hemagglutinin) to neutralize binding of the viral protein (e.g., hemagglutinin protein) to a host cell (see, for example, Current Protocols in Immunonology, 19.11.1-19.11.32, Cottey, R., *et al.,* John Wiley & Sons, Inc (2001)). Assessment of stimulation of protective immunity can also be made by employing assays that measure the ability of antibodies to inhibit

hemagglutinin binding (see, for example, Burnett, F.M., *et al., J. exp. Biol. Med. Sci. 25*:227-233 (1947); Salk, J.E. *J. Immunol. 49*:87-98 (1944); Current Protocols in Immunology, 19.11.1-19.11.32, Cottey, R., *et al.,* John Wiley & Sons, Inc (2001)).

It is believed that inhibition of hemagglutinin binding is indicative of the ability of antibodies, formed from the compositions and by the methods of the invention, to neutralize the sialic acid binding sites of the naturally occurring viral hemagglutinin ("neutralization of HA binding") and, thereby, prevent infection of the host cell as a consequence of stimulating a protective immune response. Inhibition or neutralization of hemagglutinin binding is believed to correlate with an ability of an immune response to protect against a lethal dose of virus.

Neutralization of HA binding can be assessed by *in vitro* assays (See, for example, Current Protocols in Immunology 19.11.1-19.11.32, Cottey, R., *et al.*, Suppl. 42, John Wiley & Sons, Inc. (2001) and WHO Manual on Animal Influenza Diagnosis and Surveillance, Webster, R., *et al.*, pages 28-36, 48-54, 82-92 (2002)). Exemplary viral neutralization assays rely on the ability of serum to specifically bind and prevent replication of influenza virus in culture, such as in the Madin-Darby Canine Kidney (MDCK) cell line. Briefly, cells are cultured in 96 well plates in the presence of a previously titered virus and the cytopathic effect of the replicating virus is observed under a microscope. To test serum, serial dilutions of the serum are prepared and preincubated with the viral stock for 2 hours at 37°C prior to infecting the MDCK cells. The mixture is incubated for an additional 2 hours after which the virus/serum mixture is removed and replaced with fresh media. The cells are grown for 4 days. Wells are scored as positive for viral growth if at least about 50% of the cells are dead in at least about half of the wells for a given serum dilution. The reciprocal of the highest dilution of serum which protects at least about half of the cells from death, in at least about half of the wells, is considered the neutralization titer.

Alternatively, a micro-neutralization *in vitro* assay can be performed to assess neutralization of HA binding. For example, serum is diluted and preincubated with a known titer of virus and mixed with MDCK cells, as described above. After 2 days of incubation, cells are washed and fixed with acetone. The plates are developed as an ELISA using a monoclonal antibody to the influenza

nuclear antigen NP. A microneutralization titer is determined as the reciprocal of the highest dilution which yields less than about 50% of the anti-NP reading of the virus-only control wells.

The Hemagglutination Inhibition (HAI) assay is based on the HA antigen on the surface of the influenza virus agglutinating red blood cells (RBC) and preventing red blood cells from precipitating. Antibodies that specifically bind the sialic acid-binding regions of HA prevent agglutination allowing precipitation. The assay is performed in 96 well V bottom plates with fresh chicken RBC. A stock of viral antigen is titered so that about a 4-fold excess of antigen is present relative to the minimum amount needed to prevent precipitation. The test serum, which can be from several species including mouse, ferret, poultry or human, is heated to about 56°C to inactivate complement. Serial 2-fold dilutions of the inactivated serum are performed and mixed with the stock HA. After about 30 minutes at room temperature, the RBCs are added and the plate is incubated for about 30 to about 45 minutes. Results are scored by observations: agglutination results in cloudy wells while inhibition results in a "button" of red cells precipitated at the bottom of the well. Controls include RBC with no HA, which forms a button, and HA and RBC with no serum, which remains cloudy. The HAI titer of a particular serum sample is the reciprocal of the last dilution which prevents agglutination (i.e., forms a button). For example, if about a 1:128 dilution reads as a button but the 1:256 dilution does not, the HAI titer is about 128.

In one embodiment, the invention is a method of making a protein that stimulates a protective immune response in a subject, comprising the steps of separating a portion of a protein from a naturally occurring viral hemagglutinin to thereby make a protein portion, wherein the protein portion includes at least a portion of a globular head and at least a portion of at least one secondary structure that causes the globular head to essentially retain its tertiary structure, and wherein the protein portion lacks a membrane fusion domain, a transmembrane domain and a cytoplasmic domain. The nucleic acid sequence encoding the protein portion is transformed into a prokaryotic cell and the prokaryotic host cell is cultured to thereby make the protein that stimulates a protective immune response in a subject.

"A portion of a protein," or "protein portion," as used herein in reference to a naturally occurring viral hemagglutinin, refers to any part of the naturally occurring viral hemagglutinin that is less than the entire naturally occurring viral hemagglutinin. "Naturally occurring viral hemagglutinin," as used herein, refers to the entire viral hemagglutinin, as it occurs in nature.

The protein portion can further lack a signal sequence. The protein portion can further include a sialic acid binding site.

Portions of a protein of a naturally occurring viral hemagglutinin for use in the compositions and methods of the invention can be a portion of Orthomyxoviridae (influenza A, B, C), Paramyxovirus (parainfluenza, respiratory syncytial virus, Newcastle disease virus, Nipah, Measles, canine distemper, Sendai virus), Reoviridae (rotavirus), Parvoviridae (human parvovirus, porcine parvovirus), Orthopoxvirus (Monkeypox virus, Ectromelia virus), Flaviviridae (West Nile, Japanese Encephalitis, St. Louis, Murray Valley, Kunjin), Avipoxvirus (Chicken fowlpox), Nipah virus (Guillaume V., *et al.*, *J. Virol.*, 80:7546-54 (2006)); Canine distemper virus (Singethan K., *et al.*, *J Gen Virol.*, 87:1635-42 (2006)); Newcastle disease virus, (de Leeuw O.S., *et al.*, *J Gen Virol.*, 86:1759-69 (2005); and Melanson V.R., *et al.*, *J Virol.*, 78:13053-61 (2004); Deng R., *et al.*, *Virology*, 204:17-26; (1994)), Measles (Masse N., *et al.*, *J Virol.*, 78:9051-63 (2004)), Sendai virus (Tomasi M., *et al.*, *FEBS Lett.*, 11:56-60 (2003)), Human parainfluenza (Porotto M., *et al.*, *J Virol.*, 79:2383-92 (2005); Tsurudome M., *et al.*, *Virology*, 213:190-203 (1995); Bousse T., *et al.*, *Virology*, 209:654-7 (1995); and Takimoto T., *et al.*, *J Virol.*, 66:7597-600 (1992)).

Portions of a viral hemagglutinin ("protein portions") (e.g., an influenza A, an influenza B and an influenza C viral hemagglutinin) can include at least one member selected from the group consisting of protein portions referred to herein as "HA1-1," "HA1-2" and "HA1-3."

"HA1-1," as used herein, refers to a protein portion of a viral hemagglutinin that includes at least about one β-sandwich that includes the substrate binding site, which includes at least about two β-sheets, at least about two to about three short α-helixes, at least one small β-sheet and at least one additional small β-sandwich at the bottom of the molecule and at least about four disulfide bonds. The β-sandwich that

includes the substrate binding site of the HA1-1 includes at least about four β-strands as the top sheet and at least about three to about four β-strands as the bottom sheet. At least about one α-helix of the HA1-1 portion is located by the side of β-sandwich that includes the substrate binding site and at least about one to about two are located at the bottom of the β-sandwich that includes the substrate binding site. The small β-sandwich of the HA1-1 can include at least about two to about three β-strands in each β-sheet; or about three to about four β-strands. Exemplary HA1-1 protein portions include SEQ ID NOS: 8, 11, 14, 17, 20, 38, 40, 45, 47, 49, 179, 180, 181 and 182.

"HA1-2," as used herein, refers to a protein portion of a viral hemagglutinin that includes at least about one β–sandwich that includes the substrate binding site, at least about two to about three short α-helixes, at least about one small β–sheet at the bottom of the molecule and at least about two disulfide bonds. A β–strand in a viral hemagglutinin can include between about two to about 15 amino acids. A small β–strand can include about two amino acids; or between about two to about three amino acids; or between about two to four amino acids or between about two to about five amino acids. A small β–sheet can include between about two to about three β–strands; or between about three to about four β–strands. The β-sandwich that includes the substrate binding site of HA1-2 can further include at least about four β-strands as the top sheet and at least about three to about four β-strands as the bottom sheet. At least about one α-helix of the HA1-2 portion is located by the side of the β-sandwich that includes the substrate binding site and at least about one to about two are located at the bottom of the β-sandwich that includes the substrate binding site. Exemplary HA1-2 protein portions include SEQ ID NOS: 9, 12, 15, 18, 21, 24, 26, 28, 30, 32, 39, 41, 46, 48 and 50.

"HA1-3," as used herein, refers to a protein portion of a viral hemagglutinin that includes at least one β–sandwich that includes the substrate binding site, at least about two short α-helixes and at least one disulfide bond. "ß-sandwich," as used herein, refers to at least about two sets of beta-sheets that form at least about one interactive layer. "Substrate binding site," as used herein in reference to the β–sandwich, means any part of the portion of the naturally occurring viral

hemagglutinin that has the capacity to interact or bind to a molecule. For example, the β–sandwich that includes the substrate binding site of the portion can include a portion that binds sialic acid. The β-sandwich that includes the substrate binding site of HA1-3 can further include at least about four β-strands as the top sheet and at least about three β strands as the bottom sheet. At least about one α-helix of the HA1-1 portion is located by the side of the β– sandwich that includes the substrate binding site and at least one other α-helix is located at the bottom of the β–sandwich that includes the substrate binding site. A short α-helix can include less than about 5 turns (2, 3, 4, 5 turns) in an α-helix. An α-helix in a viral hemagglutinin can be between one to about 15 turns; or between about two to 15 turns. Exemplary HA1-3 protein portions include SEQ ID NOS: 10, 13, 16, 19, 22, 25, 27, 29, 31 and 33.

"A sialic acid binding site," as that phrase is used herein in reference to the portion of the protein from the naturally occurring viral hemagglutinin, means a part of the protein portion that has the capacity to interact with sialic acid residues. "A sialic acid binding site" is also referred to herein as "a sialic acid binding domain."

"At least a portion," as used herein, refers to any part of a component (*e.g.*, a globular head, a secondary structure) or molecule (*e.g.*, a protein, antigen, Toll-like Receptor, a peptide, flagellin, HA, matrix 2 protein (M2), matrix 2 ectodomain (M2e)); or the entirety of the component or the molecule. "At least a portion," is also referred to herein as a "fragment."

"At least a portion," as used herein in reference to a flagellin (e.g., fljB/STF2, *E. coli* fliC, *S. muenchen* fliC), refers to any part of the flagellin (e.g., motif C; motif N; domain 1, 2, 3) or the entirety of the flagellin that can initiate an intracellular signal transduction pathway for a Toll-like Receptor.

A single polypeptide can exhibit several types of secondary structure. Without any stabilizing interactions, a polypeptide can assume a random-coil conformation. However, secondary structures, such as alpha(α)-helices and beta(ß)-strands, can stabilize a protein or a portion of a protein. Lateral association of β-strands form β-sheets (also referred to herein as "β-pleated sheets"). Secondary structures can be located at the surfaces of the portion, the protein, or the naturally occurring protein (e.g., viral hemagglutinin, flagellin, M2e). A tertiary structure of a

protein is the three-dimensional arrangement of amino acid residues. In contrast to secondary structure, which is stabilized by, for example, hydrogen bonds, α-helices, ß-strands, tertiary structure results from hydrophobic interactions between non-polar side chains of the portion, protein or naturally occurring viral hemagglutinin. The hydrophobic interactions hold the helices strands in random coils in a compact internal scaffold. The size and shape of a protein can depend on its primary amino acid sequence, as well as the number, size and arrangement of secondary structures.

"A globular head," as that phrase is used herein, refers to a portion of a protein of a naturally occurring viral hemagglutinin that includes the receptor or sialic acid binding regions. "Globular head," is also referred to herein as a "globular domain." The globular head of viral hemagglutinin proteins has been determined based on x-ray crystallography as described, for example, by Wilson I.A., *et al. Nature 289*:366-373 (1981); Chen, J., *et al., Cell 95*:409-417 (1998); Ha Y., *et al., The EMBO Journal 21*:865-875 (2002); Russell, R.J., *et al., Virology 325*:287-296 (2004); and Cox, N.J., *et al.*, In: Toply and Wilson's Microbiology and Microbial Infections, eds. BWJ Mathy, et al., Vol. 1 (9th ed.) New York, NY, Oxford Univ. Press, Ch. 32, p. 634 (1998). The globular head of a naturally occurring viral hemagglutinin is a component of the HA1 subunit of, for example, influenza viral hemagglutinin. In addition to the receptor binding domain, the globular head can include the E subdomain and F subdomain as described, for example, by Ha, Y., *et al.* The *EMBO Journal 21*:865-875 (2002).

The phrase, "causes the globular head to essentially retain its tertiary structure," as used herein, refers to maintenance of the tertiary structure of the globular head of the naturally occurring viral hemagglutinin sufficient to stimulate a protective immune response in a subject.

The membrane fusion domain of a viral hemagglutinin is that region of the viral hemagglutinin (involved in binding of the viral hemagglutinin) that binds a host cell. A transmembrane domain of the viral hemagglutinin is that portion of the viral hemagglutinin that spans the membrane of the virus. A cytoplasmic domain of a viral hemagglutinin is that portion of the viral hemagglutinin located on the cytoplasmic surface of the virus.

The portion of the protein of the naturally occurring viral hemagglutinin (also referred to herein as "protein portion") can further lack a signal sequence. The portion of a globular head employed in the methods described herein can include at least a portion of at least one secondary structure that includes at least a portion of at least one ß-pleated sheet; at least one alpha helix and/or at least one member selected from the group consisting of a salt bridge, a leucine zipper and a zinc finger. The portion of a globular head can further include at least about one disulfide bond, at least about two disulfide bonds, at least about three disulfide bonds, at least about four disulfide bonds, at least about five disulfide bonds and at least about six disulfide bonds.

The method of making a protein that stimulates a protective immune response in a subject can further include the step of substituting a nucleic acid sequence encoding at least one amino acid residue selected from the group consisting of a hydrophilic amino acid residue, a polar amino acid residue and a neutral amino acid residue for a nucleic acid sequence that encodes at least one hydrophobic amino acid residue in the protein portion. The hydrophobic amino acid residue substituted can include at least one member selected from the group consisting of a phenylalanine residue, a tryptophan residue and tyrosine residue. The polar amino acid residue substituted for the hydrophobic amino acid can include at least one member selected from the group consisting of an aspartic acid residue and a glutamic acid residue.

The portion of a protein of a naturally occurring viral hemagglutinin can be a portion of a naturally occurring influenza viral hemagglutinin protein (e.g., influenza A, B and C). The influenza A viral hemagglutinin protein can be at least one member selected from the group consisting of H1, H2, H3, H5, H7 and H9.

The host cell employed in the methods described herein can be a prokaryotic host cell. The prokaryotic host cell can be at least one member selected from the group consisting of an *E. coli* prokaryotic host cell, a *Pseudomonas* prokaryotic host cell, a *Bacillus* prokaryotic host cell, a *Salmonella* prokaryotic host cell and a *P. fluorescens* prokaryotic host cell.

The method of making a protein that stimulates a protective immune response in a subject can further include the step of transforming the prokaryotic

host cell with a chaperone nucleic acid sequence. The chaperone nucleic acid sequence can be at least one member selected from the group consisting of a groES-groEL chaperone, a dnaK-dnaJ-grpE chaperone, a groES-groEL-tig chaperone and a tig chaperone.

The method of making a protein that stimulates a protective immune response in a subject can further include the step of fusing at least a portion of a Toll-like Receptor (TLR) agonist to the protein. The nucleic acid sequence encoding at least a portion of a Toll-like Receptor agonist can be operably linked to the nucleic acid sequence encoding the protein portion of the viral hemagglutinin. A linker (e.g., peptide linker) can be between the Toll-like Receptor agonist and the portion of the viral hemagglutinin.

Toll-like Receptors refer to a family of receptor proteins that are homologous to the *Drosophila melangogaster* Toll protein. Toll-like Receptors are type I transmembrane signaling receptor proteins characterized by an extracellular leucine-rich repeat domain and an intracellular domain homologous to an interleukin 1 receptor. Toll-like Receptors include TLR1, TLR2, TLR3, TLR4, TLR5, TLR6, TLR7, TLR 8, TLR9, TLR10, TLR11 and TLR12.

"Agonist," as used herein in referring to a TLR, means a molecule that activates a TLR signaling pathway. A TLR signaling pathway is an intracellular signal transduction pathway employed by a particular TLR that can be activated by a TLR ligand or a TLR agonist. Common intracellular pathways are employed by TLRs and include, for example, NF-κB, Jun N-terminal kinase and mitogen-activated protein kinase. The Toll-like Receptor agonist can include at least one member selected from the group consisting of a TLR1 agonist, a TLR2 agonist (e.g., Pam3Cys, Pam2Cys, bacterial lipoprotein), a TLR3 agonist (e.g., dsRNA), a TLR4 agonist (e.g., bacterial lipopolysaccharide), a TLR5 agonist (e.g., a flagellin), a TLR6 agonist, a TLR7 agonist, a TLR8 agonist, a TLR9 agonist (e.g., unmethylated DNA motifs), TLR10 agonist, a TLR11 agonist and a TLR12 agonist.

The Toll-like Receptor agonists for use in the methods and compositions of the invention can be a Toll-like Receptor agonist component that is at least a portion of a Toll-like Receptor agonist, wherein the Toll-like Receptor agonist component includes at least one cysteine residue in a position where a cysteine does not occur in

the native Toll-like Receptor agonist, whereby the Toll-like Receptor agonist component activates a Toll-like Receptor.

TLR4 ligands (e.g., TLR4 agonists) for use in the compositions and methods of the invention can include at least one member selected from the group consisting of SEQ ID NOS: 359-406 (see, PCT/US 2006/002906/WO 2006/083706; PCT/US 2006/003285/WO 2006/083792; PCT/US 2006/041865; PCT/US 2006/042051).

GGKSGRTG	(SEQ ID NO: 359)
KGYDWLVVG	(SEQ ID NO: 360)
EDMVYRIGVP	(SEQ ID NO: 361)
VKLSGS	(SEQ ID NO: 362)
GMLSLALF	(SEQ ID NO: 363)
CVVGSVR	(SEQ ID NO: 364)
IVRGCLGW	(SEQ ID NO: 365)
AAEERTLG	(SEQ ID NO: 366)
WARVVGWLR	(SEQ ID NO: 367)
SEGYRLFGG	(SEQ ID NO: 368)
LVGGVVRRGS	(SEQ ID NO: 369)
GRVNDLWLAA	(SEQ ID NO: 370)
SGWMLWREGS	(SEQ ID NO: 371)
ERMEDRGGDL	(SEQ ID NO: 372)
KLCCFTECM	(SEQ ID NO: 373)
AVGSMERGRG	(SEQ ID NO: 374)
RDWVGGDLV	(SEQ ID NO: 375)
FFEVAKISQQ	(SEQ ID NO: 376)
WWYWC	(SEQ ID NO: 377)
MHLCSHA	(SEQ ID NO: 378)
WLFRRIG	(SEQ ID NO: 379)
YWFWRIG	(SEQ ID NO: 380)
MHLYCIA	(SEQ ID NO: 381)
WPLFPWIV	(SEQ ID NO: 382)
DMRSHAR	(SEQ ID NO: 383)
MHLCTHA	(SEQ ID NO: 384)
NLFPFY	(SEQ ID NO: 385)
MHLCTRA	(SEQ ID NO: 386)
RHLWYHA	(SEQ ID NO: 387)
WPFSAYW	(SEQ ID NO: 388)
WYLRGS	(SEQ ID NO: 389)
GKGTDLG	(SEQ ID NO: 390)
IFVRMR	(SEQ ID NO: 391)
WLFRPVF	(SEQ ID NO: 392)
FLGWLMG	(SEQ ID NO: 393)
MHLWHHA	(SEQ ID NO: 394)
WWFPWKA	(SEQ ID NO: 395)
WYLPWLG	(SEQ ID NO: 396)
WPFPRTF	(SEQ ID NO: 397)
WPFPAYW	(SEQ ID NO: 398)
FLGLRWL	(SEQ ID NO: 399)
SRTDVGVLEV	(SEQ ID NO: 400)

REKVSRGDKG	(SEQ ID NO: 401)
DWDAVESEYM	(SEQ ID NO: 402)
VSSAQEVRVP	(SEQ ID NO: 403)
LTYGGLEALG	(SEQ ID NO: 404)
VEEYSSSGVS	(SEQ ID NO: 405)
VCEVSDSVMA	(SEQ ID NO: 406)

TLR2 ligands (e.g., TLR2 agonists) for use in the compositions and methods of the invention can also include at least one member selected from the group consisting of SEQ ID NOS: 455-494 (see, PCT/US 2006/002906/WO 2006/083706; PCT/US 2006/003285/WO 2006/083792; PCT/US 2006/041865; PCT/US 2006/042051).

NPPTT	(SEQ ID NO: 455)
MRRIL	(SEQ ID NO: 456)
MISS	(SEQ ID NO: 457)
RGGSK	(SEQ ID NO: 458)
RGGF	(SEQ ID NO: 459)
NRTVF	(SEQ ID NO: 460)
NRFGL	(SEQ ID NO: 461)
SRHGR	(SEQ ID NO: 462)
IMRHP	(SEQ ID NO: 463)
EVCAP	(SEQ ID NO: 464)
ACGVY	(SEQ ID NO: 465)
CGPKL	(SEQ ID NO: 466)
AGCFS	(SEQ ID NO: 467)
SGGLF	(SEQ ID NO: 468)
AVRLS	(SEQ ID NO: 469)
GGKLS	(SEQ ID NO: 470)
VSEGV	(SEQ ID NO: 471)
KCQSF	(SEQ ID NO: 472)
FCGLG	(SEQ ID NO: 473)
PESGV	(SEQ ID NO: 474)
DPDSG	(SEQ ID NO: 475)
IGRFR	(SEQ ID NO: 476)
MGTLP	(SEQ ID NO: 477)
ADTHQ	(SEQ ID NO: 478)
HLLPG	(SEQ ID NO: 479)
GPLLH	(SEQ ID NO: 480)
NYRRW	(SEQ ID NO: 481)
LRQGR	(SEQ ID NO: 482)
IMWFP	(SEQ ID NO: 483)
RVVAP	(SEQ ID NO: 484)
IHVVP	(SEQ ID NO: 485)
MFGVP	(SEQ ID NO: 486)

CVWLQ (SEQ ID NO: 487)
IYKLA (SEQ ID NO: 488)
KGWF (SEQ ID NO: 489)
KYMPH (SEQ ID NO: 490)
VGKND (SEQ ID NO: 491)
THKPK (SEQ ID NO: 492)
SHIAL (SEQ ID NO: 493)
AWAGT (SEQ ID NO: 494)

The TLR2 ligand (e.g., TLR2 agonist) can also include at least a portion of at least one member selected from the group consisting of flagellin modification protein FlmB of *Caulobacter crescentus*; Bacterial Type III secretion system protein; invasin protein of *Salmonella*; Type 4 fimbrial biogenesis protein (PilX) of *Pseudomonas; Salmonella* SciJ protein; putative integral membrane protein of *Streptomyces*; membrane protein of *Pseudomonas*; adhesin of *Bordetella pertusis*; peptidase B of *Vibrio cholerae*; virulence sensor protein of *Bordetella*; putative integral membrane protein of *Neisseria meningitidis*; fusion of flagellar biosynthesis proteins FliR and FlhB of *Clostridium*; outer membrane protein (porin) of *Acinetobacter*; flagellar biosynthesis protein FlhF of *Helicobacter*; ompA related protein of *Xanthomonas*; omp2a porin of *Brucella*; putative porin/fimbrial assembly protein (LHrE) of *Salmonella*; wbdk of *Salmonella*; Glycosyltransferase involved in LPS biosynthesis; *Salmonella* putative permease.

The TLR2 ligand (e.g., TLR agonist) can include at least a portion of at least one member selected from the group consisting of lipoprotein/lipopeptides (a variety of pathogens); peptidoglycan (Gram-positive bacteria); lipoteichoic acid (Gram-positive bacteria); lipoarabinomannan (mycobacteria); a phenol-soluble modulin (*Staphylococcus epidermidis*); glycoinositolphospholipids (*Trypanosoma Cruzi*); glycolipids (*Treponema maltophilum*); porins (*Neisseria*); zymosan (fungi) and atypical LPS (*Leptospira interrogans* and *Porphyromonas gingivalis*).

The TLR2 ligand (e.g., TLR2 agonist) can also include at least one member selected from the group consisting of SEQ ID NOS: 495-497 (see, PCT/US 2006/002906/WO 2006/083706; PCT/US 2006/003285/WO 2006/083792; PCT/US 2006/041865; PCT/US 2006/042051).

KGGVGPVRRSSRLRRTTQPG (SEQ ID NO: 495)
GRRGLCRGCRTRGRIKQLQSAHK (SEQ ID NO: 496)
RWGYHLRDRKYKGVRSHKGVPR (SEQ ID NO: 497)

The TLR2 agonist can include at least a portion of a bacterial lipoprotein (BLP).

The TLR2 agonist can be a bacterial lipoprotein, such as Pam2Cys (S-[2,3-bis(palmitoyloxy) propyl] cysteine), Pam3Cys ([Palmitoyl]-Cys((RS)-2,3-di(palmitoyloxy)-propyl cysteine) or *Pseudomonas aeruginosa* OprI lipoprotein (OprI). Exemplary OprI lipoproteins include SEQ ID NO: 782, encoded by SEQ ID NO: 783. An exemplary *E. coli* bacterial lipoprotein for use in the invention described herein is SEQ ID NO: 784 encoded by SEQ ID NO: 785. A bacterial lipoprotein that activates a TLR2 signaling pathway (a TLR2 agonist) is a bacterial protein that includes a palmitoleic acid (Omueti, K.O., *et al., J. Biol. Chem. 280:* 36616-36625 (2005)). For example, expression of SEQ ID NOS: 783 and 785 in bacterial expression systems (e.g., *E. coli*) results in the addition of a palmitoleic acid moiety to a cysteine residue of the resulting protein (e.g., SEQ ID NOS: 782 and 784) thereby generating a TLR2 agonist for use in the compositions, fusion proteins and polypeptides of the invention. Production of tripalmitoylated-lipoproteins (also referred to as triacyl-lipoproteins) in bacteria occurs through the addition of a diacylglycerol group to the sulfhydryl group of a cysteine (e.g., cysteine 21 of SEQ ID NO: 784) followed by cleavage of the signal sequence and addition of a third acyl chain to the free N-terminal group of the same cysteine (e.g., cysteine 21 of SEQ ID NO: 784) (Sankaran, K., *et al., J. Biol. Chem. 269:*19706 (1994)), to generate a tripalmitylated peptide (a TLR2 agonist) as shown, for example, in Figure 74.

The Toll-like Receptor agonist in the compositions of the invention can further include at least one cysteine residue at the terminal amino acid of the amino-terminus and/or the terminal amino acid of the carboxy-terminus of the Toll-like Receptor agonist. For example, SEQ ID NO: 359 can further include at least one cysteine residue in a peptide bond to the amino-terminal glycine residue and/or at least one cysteine residue in a peptide bond to the carboxy-terminal glycine residue; SEQ ID NO: 360 can further include at least one cysteine residue in a peptide bond to the amino-terminal lysine residue and/or at least one cysteine residue in a peptide bond to the carboxy-terminal glycine residue; SEQ ID NO: 361 can further include at least one cysteine residue in a peptide bond to the amino-terminal glutamic acid

residue and/or at least one cysteine residue in a peptide bond to the carboxy-terminal proline residue.

TLR5 agonists for use in the methods of the invention can include at least a portion of a flagellin. Flagellin are pathogen-associated molecular patterns (PAMPs) that activate TLR5.

The flagellin in the compositions and methods described herein can be at least a portion of a *S. typhimurium flagellin* (Genbank Accession Number AF045151); at least a portion of the *S. typhimurium* flagellin selected from the group consisting of SEQ ID NO: 498, SEQ ID NO: 812, SEQ ID NO: 816 and SEQ ID NO: 500; at least a portion of an *S. muenchen* flagellin (Genbank Accession Number AB028476) that includes at least a portion of SEQ ID NO: 504 and SEQ ID NO: 813; at least a portion of *P. aeruginosa* flagellin that includes at least a portion of SEQ ID NO: 815; at least a portion of a *Listeria monocytogenes* flagellin that includes at least a portion of SEQ ID NO: 820; at least a portion of an *E. coli* flagellin that includes at least a portion of SEQ ID NO: 502 and SEQ ID NO: 814; at least a portion of a *Yersinia* flagellin; and at least a portion of a *Campylobacter* flagellin.

The flagellin employed in the compositions of the invention can also include the polypeptides of SEQ ID NO: 498, SEQ ID NO: 500, SEQ ID NO: 504 and SEQ ID NO: 502; at least a portion of SEQ ID NO: 498, at least a portion of SEQ ID NO: 500, at least a portion of SEQ ID NO: 504 and at least a portion of SEQ ID NO: 502; and a polypeptide encoded by SEQ ID NO: 499, SEQ ID NO: 501, SEQ ID NO: 505 and SEQ ID NO: 503; or at least a portion of a polypeptide encoded by SEQ ID NO: 499, SEQ ID NO: 501, SEQ ID NO: 505 and SEQ ID NO: 503.

The flagellin employed in the compositions and method of the invention can lack at least a portion of a hinge region. Hinge regions are the hypervariable regions of a flagellin. Hinge regions of a flagellin are also referred to herein as "D3 domain or region," "propellor domain or region," "hypervariable domain or region" and "variable domain or region." "Lack" of a hinge region of a flagellin, means that at least one amino acid or at least one nucleic acid codon encoding at least one amino acid that comprises the hinge region of a flagellin is absent in the flagellin. Examples of hinge regions include amino acids 176-415 of SEQ ID NO: 498, which

are encoded by nucleic acids 528-1245 of SEQ ID NO: 499; amino acids 174-422 of SEQ ID NO: 502, which are encoded by nucleic acids 522-1266 of SEQ ID NO: 503; or amino acids 173-464 of SEQ ID NO: 504, which are encoded by nucleic acids 519-1392 of SEQ ID NO: 505. Thus, if amino acids 176-415 were absent from the flagellin of SEQ ID NO: 498, the flagellin would lack a hinge region. A flagellin lacking at least a portion of a hinge region is also referred to herein as a "truncated version" of a flagellin.

"At least a portion of a hinge region," as used herein, refers to any part of the hinge region of the flagellin, or the entirety of the hinge region. "At least a portion of a hinge region" is also referred to herein as a "fragment of a hinge region." At least a portion of the hinge region of fljB/STF2 can be, for example, amino acids 200-300 of SEQ ID NO: 498. Thus, if amino acids 200-300 were absent from SEQ ID NO: 498, the resulting amino acid sequence of STF2 would lack at least a portion of a hinge region.

Alternatively, at least a portion of a naturally occurring flagellin can be replaced with at least a portion of an artificial hinge region. "Naturally occurring," in reference to a flagellin amino acid sequence, means the amino acid sequence present in the native flagellin (e.g., *S. typhimurium* flagellin, *S. muenchin* flagellin, *E. coli* flagellin). The naturally occurring hinge region is the hinge region that is present in the native flagellin. For example, amino acids 176-415 of SEQ ID NO: 498, amino acids 174-422 of SEQ ID NO: 502 and amino acids 173-464 of SEQ ID NO: 504, are the amino acids corresponding to the natural hinge region of STF2, *E. coli* fliC and *S. muenchen* flagellins, fliC, respectively. "Artificial," as used herein in reference to a hinge region of a flagellin, means a hinge region that is inserted in the native flagellin in any region of the flagellin that contains or contained the native hinge region.

The hinge region of a flagellin be deleted and replaced with an antigen (e.g., HA1-1, HA1-2, the maturational cleavage site) and the resulting construct fused to another antigen (e.g., 4xM2e).

An artificial hinge region may be employed in a flagellin that lacks at least a portion of a hinge region, which may facilitate interaction of the carboxy-and amino-terminus of the flagellin for binding to TLR5 and, thus, activation of the TLR5

innate signal transduction pathway. A flagellin lacking at least a portion of a hinge region is designated by the name of the flagellin followed by a "Δ." For example, an STF2 (e.g., SEQ ID NO: 498) that lacks at least a portion of a hinge region is referenced to as "STF2Δ" or "fljB/ STF2Δ" (e.g., SEQ ID NO: 500).

The flagellin for use in the methods and compositions of the invention can be a at least a portion of a flagellin, wherein the flagellin component includes at least one cysteine residue and whereby the flagellin component activates a Toll-like Receptor 5; a flagellin component that is at least a portion of a flagellin, wherein at least one lysine of the flagellin component has been substituted with at least one arginine, whereby the flagellin component activates a Toll-like Receptor 5; a flagellin component that is at least a portion of a flagellin, wherein at least one lysine of the flagellin component has been substituted with at least one serine residue, whereby the flagellin component activates a Toll-like Receptor 5; a flagellin component that is at least a portion of a flagellin, wherein at least one lysine of the flagellin component has been substituted with at least one histidine residue, whereby the flagellin component activates a Toll-like Receptor 5, as described herein.

A recombinant fusion protein can be generated by operably linking a TLR agonist to the protein portion (e.g., HA1-1, HA1-2).

"Fusion protein," as used herein, refers to a protein generated from at least two similar or distinct components (e.g., a protein portion of HA and a TLR agonist). Fusion proteins can be generated recombinantly or by chemical conjugation.

In an embodiment, fusion proteins of protein portions of HA and TLR agonists (e.g., SEQ ID NOS: 89-92, 95, 151-160, 177, 209, 210 and 211) can be admixed or coadministered with fusion proteins of TLR agonists and M2e proteins (e.g., SEQ ID NOS: 528, 587, 589 and 591) and administered to a subject to stimulate an immune response, such as a protective immune response.

Fusion proteins of the invention can be designated by components of the fusion proteins separated by a ".". For example, "STF2.HA1-2" refers to a fusion protein comprising one STF2 protein and one HA1-2 protein; and "STF2Δ.HA1-2" refers to a fusion protein comprising one STF2 protein without the hinge region and

HA1-2. Exemplary fusion proteins of the invention include SEQ ID NOS: 89-92, 95, 151-160, 177, 209, 210 and 211.

The fusion proteins can include, for example, two, three, four, five, six or more Toll-like Receptor agonists (e.g., flagellin) and two, three, four, five, six or more antigens (e.g., protein portions, such as HA1-1, HA1-2). When two or more TLR agonists and/or two or more protein portions comprise fusion proteins of the invention, they are also referred to as "multimers." For example, a multimer of HA1-1 can be four HA1-1 sequences, which is referred to herein as 4xHA1-1. Likewise, "2xHA1-1" is a multimer of two HA1-1 sequences. (See, for example, SEQ ID NOS: 342-346 and 348).

The fusion proteins of the invention can further include a linker between at least one component of the fusion protein (e.g., TLR agonist) and at least one other component of the fusion protein (e.g., HA1-1, HA1-2) of the composition, a linker between at least two of similar components of the fusion protein (e.g., HA1-1, HA1-2) or any combination thereof. "Linker," as used herein in reference to a fusion protein of the invention, refers to a connector between components of the fusion protein in a manner that the components of the fusion protein are not directly joined. For example, one part of the fusion protein (e.g., flagellin) can be linked to a distinct part (e.g., protein portion, HA1-1, HA1-2, an antigen) of the fusion protein. Likewise, at least two or more similar or like components of the fusion protein can be linked (e.g., two flagellin can further include a linker between each flagellin, or two HA proteins can further include a linker between each HA protein).

Additionally, or alternatively, the fusion proteins of the invention can include a combination of a linker between distinct components of the fusion protein and similar or like components of the fusion protein. For example, a fusion protein can comprise at least two TLR agonists that further includes a linker between, for example, two or more flagellin; at least two protein portions of HA that further include a linker between them; a linker between one component of the fusion protein (e.g., flagellin) and another distinct component of the fusion protein (e.g., protein portions of HA), or any combination thereof.

The linker can be an amino acid linker. The amino acid linker can include synthetic or naturally occurring amino acid residues. The amino acid linker

employed in the fusion proteins of the invention can include at least one member selected from the group consisting of a lysine residue, a glutamic acid residue, a serine residue and an arginine residue. The amino acid linker can include, for example, SEQ ID NOS: 521, 523, 524 and 526, encoded by the nucleic acid sequences of SEQ ID NOS: 520, 522, 525 and 527 respectively.

The Toll-like Receptor agonist of the fusion proteins of the invention can be fused to a carboxy-terminus, the amino-terminus or both the carboxy- and amino-terminus of the protein portion of HA (e.g., HA1-1, HA1-2 or other antigens, such as the ectodomain of the Matrix 2 protein, a maturational cleavage site) .

Fusion proteins can be generated by fusing the protein portions of HA (or other antigens, such as the maturational cleavage site of HA) to at least one of four regions (Regions 1, 2, 3 and 4) of flagellin, which have been identified based on the crystal structure of flagellin (PDB:1UCU) (see Figure 77).

Region 1 is TIAL (SEQ ID NO: 823)...-...GLG (95-111 of SEQ ID NO: 841). The corresponding residues for *Salmonella typhimurium* fljB construct are TTLD (SEQ ID NO: 824)...-...GTN (196-216 of SEQ ID NO: 841). This region is an extended peptide sitting in a groove of two beta strands (GTDQKID (SEQ ID NO: 825) and NGEVTL (SEQ ID NO: 826) of (SEQ ID NO: 841). Substitution of this peptide with an antigen (e.g., HA maturational cleavage site) may mimic the conformation of wild type maturational cleavage site peptide in HA0. Exemplary amino acids that may be substituted include: flagellin residues SGLDDAAIKAAT (SEQ ID NO: 827) (201-212 of SEQ ID NO: 841) substituted with maturational cleavage site residues: RGIFGAIAGFIE (SEQ ID NO: 828), which correspond to the A/H3N2 subtype, or RGLFGAIAGFIE (SEQ ID NO: 803), which correspond to the maturational cleavage site residues from the A/H2N1, A/H1N1, A/H5N1 subtypes or with RGFFGAIAGFLE (SEQ ID NO: 805), which correspond to Influenza B HA maturational cleavage site residues.

Region 2 of the *Salmonella* flagellin is a small loop GTG (238-240 of SEQ ID NO: 841) in 1UCU structure (see Figure 77). The corresponding loop in *Salmonella* fljB is GADAA (SEQ ID NO: 829) (244-248 of SEQ ID NO: 841). Insertion of an antigen (e.g., protein portions of HA, a maturation cleavage site peptide) in this loop or replacement of the entire loop with an antigen (e.g., protein

portions of HA, a maturational cleavage site peptide) should preserve the extended loop structure of the maturational cleavage site peptide that is associated with the native HA molecule.

Region 3 is a bigger loop that resides on the opposite side of the Region 1 peptide (see Figure 77). This loop can be simultaneously substituted together with region 1 to create a double copy of the antigen (e.g., protein portions of HA, a maturational cleavage site peptide). The loop starts from AGGA (SEQ ID NO: 830) and ends at PATA (SEQ ID NO: 831) (259-274 of SEQ ID NO: 841). The corresponding *Salmonella* fljB sequence is AAGA (SEQ ID NO: 832...-...ATTK (SEQ ID NO: 833) (266-281 of SEQ ID NO: 841). The sequence AGATKTTMPAGA (SEQ ID NO: 834) (267-278 of SEQ ID NO: 841) can be replaced with the antigens (e.g., protein portions of HA, a maturation cleavage site peptides).

Region 4 is the loop (VTGTG (SEQ ID NO: 835)) connecting a short α-helix (TEAKAALTAA (SEQ ID NO: 836)) and a β-strand (ASVVKMSYTDN (SEQ ID NO: 837) in 1UCU structure (see Figure 77). The corresponding loop in *Salmonella* fljB is a longer loop VDATDANGA (SEQ ID NO: 838 (307-315 of SEQ ID NO: 841). An antigen (e.g., a protein portion of HA, a maturation cleavage site peptide) can be inserted into or replace this region.

One or more copies of antigens (e.g., maturation cleavage site) can be inserted or used to replace the peptides listed in the above four regions. Preferably, the replacements would be in Region 1 and Region 3.

Exemplary sequences of maturation cleavage site peptides of HA are listed below:

Sequence	Subtype
NVPEKQTRGIFGAIAGFIE	A/H3N2 (SEQ ID NO: 800)
NVPQIESRGLFGAIAGFIE	A/H2N1 (SEQ ID NO: 801)
NIPSIQSRGLFGAIAGFIE	A/H1N1 (SEQ ID NO: 802)
RERRRKKRGLFGAIAGFIE	A/H5N1 (SEQ ID NO: 839)
PAKLLKERGFFGAIAGFLE	B/HA (SEQ ID NO: 804)

```
Sequence alignment of X-ray model (1UCU) and SEQ ID NO: 841

75.1% identity in 506 residues overlap; Score: 1703.0; Gap frequency: 2.6%

1UCU   1   AQVINTNSLSLLTQNNLNKSQSALGTAIERLSSGLRINSAKDDAAGQAIANRFTANIKGL
Sfla   2   AQVINTNSLSLLTQNNLNKSQSALGTAIERLSSGLRINSAKDDAAGQAIANRFTANIKGL
           ************************************************************

1UCU   61  TQASRNANDGISIAQTTEGALNEINNNLQRVRELAVQSANSTNSQSDLDSIQAEITQRLN
Sfla   62  TQASRNANDGISIAQTTEGALNEINNNLQRVRELAVQSANSTNSQSDLDSIQAEITQRLN
           ************************************************************

1UCU   121 EIDRVSGQTQFNGVKVLAQDNTLTIQVGANDGETIDIDLKQINSQTLGLDTLNVQQKYKV
Sfla   122 EIDRVSGQTQFNGVKVLAQDNTLTIQVGANDGETIDIDLKQINSQTLGLDSLNVQKAYDV
           ************************************************** **** * *

1UCU   181 SDTAATVTGYAD--TTI---ALDNSTFKASATGLGGTDQKIDGDLKFDDTTGKYYAKVTV
Sfla   182 KDTAVTTKAYANNGTTLDVSGLDDAAIKAATGGTNGTASVTGGAVKFDADNNKYFVTIGG
            *** *   **   **      **    **  *  **     *  ***    **

1UCU   236 TGGT—GKDGYYEVSVDKTNGEVTLAGGATSPLTGGLPATATEDVKNVQVANADLTEAKA
Sfla   242 FTGADAAKNGDYEVNV-ATDGTVTLAAGATKTTMPAGATTKTEVQELKDTPAVVSADAKN
             *    * * *** *  * * **** ***         * **                **

1UCU   294 ALTAAGVTGT----ASVVKMSYTDNNGKTIDGGLAVKVGDDYYSATQNK-DGSISINTTK
Sfla   301 ALIAGGVDATDANGAELVKMSYTDKNGKTIEGGYALKAGDKYYAADYDEATGAIKAKTTS
           ** * **  *    *  ******* ***** ** * * ** ** *      * *   **

1UCU   349 YTADDGTSKTALNKLGGADGKTEVVSIGGKTYAASKAEGHNFKAQPDLAEAAATTTENPL
Sfla   361 YTAADGTTKTAANQLGGVDGKTEVVTIDGKTYNASKAAGHDFKAQPELAEAAAKTTENPL
           *** *** *** * *** ******* * **** **** ** ***** ****** ******

1UCU   409 QKIDAALAQVDTLRSDLGAVQNRFNSAITNLGNTVNNLTSVRSRIEDSDYATEVSNMSRA
Sfla   421 QKIDAALAQVDALRSDLGAVQNRFNSAITNLGNTVNNLSEARSRIEDSDYATEVSNMSRA
           *********** ************************   ******************

1UCU   469 QILQQAGTSVLAQANQVPQNVLSLLR (SEQ ID NO: 840)
Sfla   481 QILQQAGTSVLAQANQVPQNVLSLLR (SEQ ID NO: 841)
           **************************
```

The methods of making a protein that stimulates a protective immune response in a subject can further include the step of operably linking a nucleic acid sequence encoding a carrier protein to the nucleic acid sequence encoding a portion of the viral hemagglutinin.

"Carrier," as used herein, refers to a molecule (e.g., protein, peptide) that can enhance stimulation of a protective immune response. Carriers can be physically attached (e.g., linked by recombinant technology, peptide synthesis, chemical conjugation or chemical reaction) to a composition (e.g., a protein portion of a naturally occurring viral hemagglutinin) or admixed with the composition.

Carriers for use in the methods and compositions described herein can include, for example, at least one member selected from the group consisting of Tetanus toxoid (TT), *Vibrio cholerae* toxoid, Diphtheria toxoid (DT), a cross-

reactive mutant (CRM) of diphtheria toxoid, *E. coli* enterotoxin, *E. coli* B subunit of heat labile enterotoxin (LTB), Tobacco mosaic virus (TMV) coat protein, protein Rabies virus (RV) envelope protein (glycoprotein), thyroglobulin (Thy), heat shock protein HSP 60 Kda, Keyhole limpet hemocyamin (KLH), an early secreted antigen tuberculosis-6 (ESAT-6), exotoxin A, choleragenoid, hepatitis B core antigen, and the outer membrane protein complex of *N. meningiditis* (OMPC) (see, for example, Schneerson, R., *et al., Prog Clin Biol Res* 47:77-94 (1980); Schneerson, R., *et al., J Exp Med* 152:361-76 (1980); Chu, C., *et al., Infect Immun* 40: 245-56 (1983); Anderson, P., *Infect Immun 39*:233-238 (1983); Anderson, P., *et al., J Clin Invest 76*:52-59 (1985); Fenwick, B.W., *et al., 54*:583-586 (1986); Que, J.U., *et al. Infect Immun 56*:2645-9 (1988); Que, J.U., *et al. Infect Immun 56*:2645-9 (1988); (Que, J.U., *et al. Infect Immun 56*:2645-9 (1988); Murray, K., *et al., Biol Chem 380*:277-283 (1999); Fingerut, E., *et al., Vet Immunol Immunopathol 112*:253-263 (2006); and Granoff, D.M., *et al., Vaccine 11*:Suppl 1:S46-51 (1993)).

Exemplary carrier proteins for use in the methods and compositions described herein can include at least one member selected from the group consisting of SEQ ID NOS: 788-795:

Cross-reactive mutant (CRM) of diphtheria toxin including, CRM197
GADDVVDSSKSFVMENFSSYHGTKPGYVDSIQKGIQKPKSGTQGNYDDDW
KGFYSTDNKYDAAGYSVDNENPLSGKAGGVVKVTYPGLTKVLALKVDNA
ETIKKELGLSLTEPLMEQVGTEEFIKRFGDGASRVVLSLPFAEGSSSVEYINN
WEQAKALSVELEINFETRGKRGQDAMYEYMAQACAGNRVRRSVGSSLSCI
NLDWDVIRDKTKTKIESLKEHGPIKNKMSESPNKTVSEEKAKQYLEEFHQTA
LEHPELSELKTVTGTNPVFAGANYAAWAVNVAQVIDSETADNLEKTTAALS
ILPGIGSVMGIADGAVHHNTEEIVAQSIALSSLMVAQAIPLVGELVDIGFAAY
NFVESIINLFQVVHNSYNRPAYSPGHKTQPFLHDGYAVSWNTVEDSIIRTGF
QGESGHDIKITAENTPLPIAGVLLPTIPGKLDVNKSKTHISVNGRKIRMRCRAI
DGDVTFCRPKSPVYVGNGVHANLHVAFHRSSSEKIHSNEISSDSIGVLGYQK
TVDHTKVNSKLSLFFEIKS (SEQ ID NO: 788)

Coat protein of Tobacco mosaic virus (TMV) coat protein
MMAYSIPTPSQLVYFTENYADYIPFVNRLINARSNSFQTQSGRDELREILIKS
QVSVVSPISRFPAEPAYYIYLRDPSISTVYTALLQSTDTRNRVIEVENSTNVTT
AEQLNAVRRTDDASTAIHNNLEQLLSLLTNGTGVFNRTSFESASGLTWLVTT
TPRTA (SEQ ID NO: 789)

Coat protein of alfalfa mosaic virus (AMV)
MSSSQKKAGGKAGKPTKRSQNYAALRKAQLPKPPALKVPVAKPTNTILPQT
GCVWQSLGTPLSLSSSNGLGARFLYSFLKDFAAPRILEEDLIFRMVFSITPSHA

GSFCLTDDVTTEDGRAVAHGNPMQEFPHGAFHANEKFGFELVFTAPTHAG
MQNQNFKHSYAVALCLDFDALPEGSRNPSYRFNEVWVERKAFPRAGPLRSL
ITVGLFDDADDLDRQ (SEQ ID NO: 790)

Coat protein of Potato virus X
MTTPANTTQATGSTTSTTTKTAGATPATTSGLFTIPDGEFFSTARAIVASNAV
ATNEDLSKIEAIWKDMKVPTDTMAQAAWDLVRHCADVGSSAQTEMIDTGP
YSNGISRARLAAAIKEVCTLRQFCMKYAPVVWNWMLTNNSPPANWQAQG
FKPEHKFAAFDFFNGVTNPAAIMPKEGLIRPPSEAEMNAAQTAAFVKITKAR
AQSNDFASLDAAVTRGRITGTTTAEAVVTLPPP (SEQ ID NO: 791)

Porins from *Neisseria sp*,e.g.,
class 1 outer membrane protein of *Neisseria meningitides*
MRKKLTALVLSALPLAAVADVSLYGEIKAGVEGRNYQLQLTEAQAANGGA
SGQVKVTKVTKAKSRIRTKISDFGSFIGFKGSEDLGEGLKAVWQLEQDVSVA
GGGATQWGNRESFIGLAGEFGTLRAGRVANQFDDASQAIDPWDSNNDVAS
QLGIFKRHDDMPVSVRYDSPEFSGFSGSVQFVPAQNSKSAYKPAYWTTVNT
GSATTTTFVPAVVGKPGSDVYYAGLNYKNGGFAGNYAFKYARHANVGRD
AFELFLLGSGSDQAKGTDPLKNHQVHRLTGGYEEGGLNLALAAQLDLSENG
DKTKNSTTEIAATASYRFGNAVPRISYAHGFDFIERGKKGENTSYDQIIAGVD
YDFSKRTSAIVSGAWLKRNTGIGNYTQINAASVGLRHKF (SEQ ID NO: 792)

Major fimbrial subunit protein type I (Fimbrillin)
MVLKTSNSNRAFGVGDDESKVAKLTVMVYNGEQQEAIKSAENATKVEDIK
CSAGQRTLVVMANTGAMELVGKTLAEVKALTTELTAENQEAAGLIMTAEP
KTIVLKAGKNYIGYSGTGEGNHIENDPLKIKRVHARMAFTEIKVQMSAAYD
NIYTFVPEKIYGLIAKKQSNLFGATLVNADANYLTGSLTTFNGAYTPANYAN
VPWLSRNYVAPAADAPQGFYVLENDYSANGGTIHPTILCVYGKLQKNGAD
LAGADLAAAQAANWVDAEGKTYYPVLVNFNSNNYTYDSNYTPKNKIERN
HKYDIKLTITGPGTNNPENPITESAHLNVQCTVAEWVLVGQNATW (SEQ ID
NO: 793)

Mycoplasma fermentans macrophage activating lipopeptide (MALP-2)
MKKSKKILLGLSPIAAVLPAVAVSCGNNDESNISFKEKDISKYTTTNANGKQ
VVKNAELLKLKPVLITDEGKIDDKSFNQSAFEALKAINKQTGIEINSVEPSSN
FESAYNSALSAGHKIWVLNGFKHQQSIKQYIDAHREELERNQIKIIGIDFDIET
EYKWFYSLQFNIKESAFTTGYAIASWLSEQDESKRVVASFGVGAFPGVTTFN
EGFAKGILYYNQKHKSSKIYHTSPVKLDSGFTAGEKMNTVINNVLSSTPADV
KYNPHVILSVAGPATFETVRLANKGQYVIGVDSDQGMIQDKDRILTSVLKHI
KQAVYETLLDLILEKEEGYKPYVVKDKKADKKWSHFGTQKEKWIGVAENH
FSNTEEQAKINNKIKEAIKMFKELPEDFVKYINSDKALKDGNKIDNVSERLE
AIISAINKAAK (SEQ ID NO: 794)

p19 protein of *Mycobacterium tuberculosis*
ATTLPVQRHPRSLFPEFSELFAAFPSFAGLRPTFDTRLMRLEDEMKEGRYEV
RAELPGVDPDKDVDIMVRDGQLTIKAERTEQKDFDGRSEFAYGSFVRTVSL
PVGADEDDIKATYDKGILTVSVAVSEGKPTEKHIQIRSTN (SEQ ID NO: 795)

The compositions of the invention can further include at least one adjuvant. Adjuvants contain agents that can enhance the immune response against substances that are poorly immunogenic on their own (see, for example, Immunology Methods Manual, vol. 2, I. Lefkovits, ed., Academic Press, San Diego, CA, 1997, ch. 13). Immunology Methods Manual is available as a four volume set, (Product Code Z37,435-0); on CD-ROM, (Product Code Z37,436-9); or both, (Product Code Z37,437-7). Adjuvants can be, for example, mixtures of natural or synthetic compounds that, when administered with compositions of the invention, such as proteins that stimulate a protective immune response made by the methods described herein, further enhance the immune response to the protein. Compositions that further include adjuvants may further increase the protective immune response stimulated by compositions of the invention by, for example, stimulating a cellular and/or a humoral response (i.e., protection from disease versus antibody production). Adjuvants can act by enhancing protein uptake and localization, extend or prolong protein release, macrophage activation, and T and B cell stimulation. Adjuvants for use in the methods and compositions described herein can be mineral salts, oil emulsions, mycobacterial products, saponins, synthetic products and cytokines. Adjuvants can be physically attached (e.g., linked by recombinant technology, by peptide synthesis or chemical reaction) to a composition described herein or admixed with the compositions described herein.

In another embodiment, the invention is a method of making a protein that stimulates a protective immune response in a subject, comprising the steps of separating a portion of a protein from a naturally occurring viral hemagglutinin to thereby form a protein portion, wherein the protein portion includes at least a portion of a globular head and at least a portion of at least one secondary structure that causes the globular head to essentially retain its tertiary structure, and wherein the protein portion lacks a membrane fusion domain, a transmembrane domain and a cytoplasmic domain; transfecting a nucleic acid sequence encoding the protein portion into a eukaryotic host cell, wherein the eukaryotic host cell is not a *Pichia pastoris* eukaryotic host cell; and culturing the eukaryotic host cell to thereby make the protein. In one embodiment, the protein portion can further lack a signal

sequence. In another embodiment, the protein portion can further include a sialic acid binding site.

In a further embodiment, the invention is a method of making a protein that stimulates a protective immune response in a subject, comprising the steps of separating a portion of a protein from a naturally occurring viral hemagglutinin to thereby form a protein portion, wherein the protein portion includes at least a portion of a globular head, and at least a portion of at least one secondary structure that causes the globular head to essentially retain its tertiary structure, and wherein the protein portion lacks a membrane fusion domain, a transmembrane domain and a cytoplasmic domain; transfecting a nucleic acid sequence encoding the protein portion into a eukaryotic host cell, wherein the eukaryotic host cell is not an insect eukaryotic host cell; and culturing the eukaryotic host cell to thereby make the protein that stimulates a protective immune response in a subject.

In a further embodiment, the invention is a method of making a protein that stimulates a protective immune response in a subject, comprising the steps of separating a portion of a protein from a naturally occurring viral hemagglutinin to thereby form a protein portion, wherein the protein portion includes at least a portion of a globular head, and at least a portion of at least one secondary structure that causes the globular head to essentially retain its tertiary structure, and wherein the protein portion lacks a membrane fusion domain, a transmembrane domain and a cytoplasmic domain; transfecting a nucleic acid sequence encoding the protein portion into a eukaryotic host cell, wherein the eukaryotic host cell is not a stably transformed insect cell; and culturing the eukaryotic host cell to thereby make the protein that stimulates a protective immune response in a subject.

In another embodiment, the invention is a method of making a protein that stimulates a protective immune response in a subject, comprising the steps of separating a portion of a protein from a naturally occurring viral hemagglutinin to thereby make a protein portion, wherein the protein portion includes at least a portion of a globular head and at least a portion of at least one secondary structure that causes the globular head to essentially retain its tertiary structure, and wherein the protein portion lacks a membrane fusion domain, a transmembrane domain and a cytoplasmic domain; infecting a nucleic acid sequence encoding the protein portion

into an insect cell host cell (e.g., a baculovirus insect host cell, such as Sf9 or High5 cells); and culturing the insect host cell to thereby make the protein that stimulates a protective immune response in a subject.

The methods of making a protein that stimulates a protective immune response in a subject can further include the step of deleting at least one glycosylation site in the nucleic acid sequence encoding the protein portion. The glycosylation site that is deleted can include an N-glycosylation site.

The eukaryotic host cells employed in the methods of the invention can include a *Saccharomyces* eukaryotic host cell, an insect eukaryotic host cell (e.g., at least one member selected from the group consisting of a *Baculovirus* infected insect cell, such as *Spodoptera frugiperda* (Sf9) or *Trichhoplusia ni* (High5) cells; and a *Drosophila* insect cell, such as Dme12 cells), a fungal eukaryotic host cell, a parasite eukaryotic host cell (e.g., a *Leishmania tarentolae* eukaryotic host cell), CHO cells, yeast cells (e.g., *Pichia*) and a *Kluyveromyces lactis* host cell.

Suitable eukaryotic host cells and vectors can also include plant cells (e.g., tomato; chloroplast; mono- and dicotyledonous plant cells; *Arabidopsis thaliana; Hordeum vulgare; Zea mays;* potato, such as *Solanum tuberosum*; carrot, such as *Daucus carota L.*; and *tobacco, such as Nicotiana tabacum, Nicotiana benthamiana* (Gils, M., *et al., Plant Biotechnol J. 3*:613-20 (2005); He, D.M., *et al., Colloids Surf B Biointerfaces*, (2006); Huang, Z., *et al., Vaccine 19*:2163-71 (2001); Khandelwal, A., *et al., Virology. 308*:207-15 (2003); Marquet-Blouin, E., *et al., Plant Mol Biol 51*:459-69 (2003); Sudarshana, M.R., *et al. Plant Biotechnol J. 4*:551-9 (2006); Varsani, A., *et al., Virus Res*, 120:91-6 (2006); Kamarajugadda S., *et al., Expert Rev Vaccines 5*:839-49 (2006); Koya V, *et al., Infect Immun. 73*:8266-74 (2005); Zhang, X., *et al., Plant Biotechnol J. 4*:419-32 (2006)).

The proteins made by the methods of the invention and the compositions of the invention can be purified and characterized employing well-known methods (e.g., gel chromatography, cation exchange chromatography, SDS-PAGE), as described herein.

For large scale production, fermentation techniques can be employed, as described herein. Additional exemplary fermentation techniques can include a proposed cycle that can start with a culture inoculated into 6 L of MRBR media, as

described herein, held at about 30°C, about pH 7, and DO controlled to greater than about 30%. A 6 liter feed can then be started at least about 30 minutes after glucose exhaustion. The proposed 6 liter feed media, when combined with 6 L of MRBR media, can provide the necessary conditions for *E. coli* growth based on about 52% utilization of carbon for growth. The feed may or may not include IPTG. The batch can be induced with at least 2 mM IPTG, introduced as a bolus, shortly after the feed is started to initiate production. The feed rate can start at about 20 mL feed per hour per liter bioreactor volume and increase over time based on the ability of the culture to accept more glucose without glucose accumulation. The culture can be harvested when the feed is complete. The 6 liter feed media, about pH 6.0, can include Glucose 180 g/L; KH_2PO_4 2 g/L; NaH_2PO_4 (H_2O) 4 g/L; $(NH_4)_2HPO_4$ 12 g/L; $(NH_4)_2HSO_4$ 4 g/L; DL-Alanine 40 g/L; Citric Acid 4 g/L; $MgSO_4(7H_20)$ 5.5 g/L; Trace Metals 6 mL; $CaCl_2$ 2.5 g/L; $FeSO_4$ $7H_2O$ 1 g/L.

Cell disruption and clarification in a large scale production can include removal of Triton X-100 from the resuspension buffer; dissolution of insolubles by the addition of 50 mM Tris, 25 mM NaCl, 8 M urea, about pH 8 to the lysate; addition of PEI (poly ethylamine) and subsequent removal by centrifugation with one or more of the buffers to remove nucleic acids and/or aid in filtration; the addition of flocullants, such as Aerosil 380, Aerosil 200, Alkoxide Alu C, and Celpur; and subsequent removal by centrifugation to aid in filtration. Cation exchange chromatography can include the use of a process resin, adding a denaturing endotoxin removal step containing up to 8 M urea and up to about 2% Triton X-100, and a step gradient elution. The step elution gradient can include about 100 to about 200 mM NaCl.

In yet another embodiment, the invention is a method of stimulating protective immunity in a subject, comprising the step of administering to the subject a composition that includes a protein made by a method comprising the steps of separating a portion of a protein from a naturally occurring viral hemagglutinin to thereby form a protein portion, wherein the protein portion includes at least a portion of a globular head and at least a portion of at least one secondary structure that causes the globular head to essentially retain its tertiary structure, and wherein the protein portion lacks a membrane fusion domain, a transmembrane domain and a

cytoplasmic domain; transforming a nucleic acid sequence encoding the portion into a prokaryotic host cell; and culturing the prokaryotic host cell to thereby make the protein that stimulates protective immunity in a subject.

In still another embodiment, the invention is a method of stimulating protective immunity in a subject, comprising the step of administering to the subject a composition that includes a protein portion of a naturally occurring viral hemagglutinin, wherein the protein portion includes at least a portion of a globular head and at least a portion of one secondary structure that causes the globular head to essentially retain its tertiary structure and wherein the protein portion lacks a membrane fusion domain, a transmembrane domain and a cytoplasmic domain.

In a further embodiment, the invention is a method of making a viral hemagglutinin protein that stimulates a protective immune response in a subject, comprising the steps of separating a portion of a protein from a naturally occurring viral hemagglutinin to thereby make a protein portion, wherein the protein portion includes at least a portion of a globular head and at least a portion of at least one secondary structure that causes the globular head to essentially retain its tertiary structure, and wherein the protein portion lacks a membrane fusion domain, a transmembrane domain and a cytoplasmic domain; transfecting a nucleic acid sequence encoding the portion in a eukaryotic host cell, wherein the eukaryotic host cell is not a *Pichia pastoris* eukaryotic host cell; and culturing the eukaryotic host cell to thereby make the protein that stimulates a protective immune response in a subject.

In yet another embodiment, the invention is a method of making a viral hemagglutinin protein that stimulates a protective immune response in a subject, comprising the steps of separating a portion of a protein from a naturally occurring viral hemagglutinin to thereby make a protein portion, wherein the protein portion includes at least a portion of a globular head and at least a portion of at least one secondary structure that causes the globular head to essentially retain its tertiary structure, and wherein the protein portion lacks a membrane fusion domain, a transmembrane domain and a cytoplasmic domain; transfecting a nucleic acid sequence encoding the portion in a eukaryotic host cell, wherein the eukaryotic host cell is not a *Drosophila melanogaster* eukaryotic host cell; and culturing the

eukaryotic host cell to thereby make the protein that stimulates a protective immune response in a subject.

In still another embodiment, the invention is a method of making a viral hemagglutinin protein that stimulates a protective immune response in a subject, comprising the steps of transforming a prokaryotic host cell with a nucleic acid sequence encoding at least one viral hemagglutinin that lacks a transmembrane domain and a cytoplasmic domain; and culturing the prokaryotic cell to thereby make the protein that stimulates a protective immune response in a subject.

An additional embodiment of the invention is a method of stimulating protective immunity in a subject, comprising the step of administering to the subject a composition that includes a protein made by a method comprising the steps of transforming a prokaryotic host cell with a nucleic acid sequence encoding at least one viral hemagglutinin that lacks a transmembrane domain and a cytoplasmic domain; and culturing the prokaryotic host cell to thereby make the protein that stimulates a protective immune response in a subject.

In another embodiment, the invention is a method of stimulating protective immunity in a subject, comprising the step of administering to the subject a composition that includes a protein having at least one viral hemagglutinin that lacks a transmembrane domain and a cytoplasmic domain, wherein the protein was expressed in a prokaryotic cell.

In still another embodiment, the invention is a composition comprising at least a portion of at least one pathogen-associated molecular pattern and a protein portion of a naturally occurring viral hemagglutinin, wherein the protein portion of the naturally occurring viral hemagglutinin includes at least a portion of a globular head and at least a portion of at least one secondary structure that causes the globular head to essentially retain its tertiary structure, and wherein the protein portion of the naturally occurring viral hemagglutinin lacks a membrane fusion domain, a transmembrane domain, and a cytoplasmic domain.

Pathogen-associated molecular patterns (PAMPs), such as a flagellin or a bacterial lipoprotein, refer to a class of molecules (e.g., protein, peptide, carbohydrate, lipid, lipopeptide, nucleic acid) found in microorganisms that, when

bound to a pattern recognition receptor (PRR), can trigger an innate immune response. The PRR can be a Toll-like Receptor (TLR).

TLRs are the best characterized type of Pattern Recognition Receptor (PRR) expressed on antigen-presenting cells (APC). APC utilize TLRs to survey the microenvironment and detect signals of pathogenic infection by engaging the cognate ligands of TLRs, PAMPs. PAMP and TLR interaction triggers the innate immune response, the first line of defense against pathogenic insult, manifested as release of cytokines, chemokines and other inflammatory mediators; recruitment of phagocytic cells; and important cellular mechanisms which lead to the expression of costimulatory molecules and efficient processing and presentation of antigens to T-cells. TLRs control both innate and the adaptive immune responses.

The binding of PAMPs to TLRs activates innate immune pathways. Target cells can result in the display of co-stimulatory molecules on the cell surface, as well as antigenic peptide in the context of major histocompatibility complex molecules (see Figure 40). The compositions, fusion proteins or polypeptides of the invention can include a PAMP (e.g., a flagellin) that binds to a TLR (e.g., TLR5), promoting differentiation and maturation of the APC, including production and display of co-stimulatory signals (see Figure 40). The compositions can be internalized by its interaction with the TLR and processed through the lysosomal pathway to generate antigenic peptides, which are displayed on the surface in the context of the major histocompatibility complex.

The compositions, fusion proteins and polypeptides of the invention employ pathogen-associated molecular patterns (TLR agonists) that trigger cellular events resulting in the expression of costimulatory molecules, secretion of critical cytokines and chemokines; and efficient processing and presentation of antigens to T-cells. As discussed above, TLRs recognize PAMPs including bacterial cell wall components (e.g., bacterial lipoproteins and lipopolysaccharides), bacterial DNA sequences that contain unmethylated CpG residues and bacterial flagellin. TLRs act as initiators of the innate immune response and gatekeepers of the adaptive immune response (Medzhitov, R., *et al., Cold Springs Harb. Symp. Quant. Biol. 64*:429 (1999); Pasare, C., *et al., Semin, Immunol 16*:23 (2004); Medzhitov, R., *et al., Nature*

388:394 (1997); Barton, G.M., *et al., Curr. Opin. Immunol 14*:380 (2002); Bendelac, A., *et al., J. Exp. Med. 195*:F19 (2002)).

As discussed above, the binding of PAMPs to TLRs activates immune pathways for use in the compositions, fusion proteins and polypeptides of the invention, which can be employed in stimulating the immune system in a subject. The compositions, fusion proteins and polypeptides of the invention can trigger an immune response to an antigen (e.g., a viral protein, such as an influenza viral) and trigger signal transduction pathways of the innate and adaptive immune system of the subject to thereby stimulate the immune system of a subject. Stimulation of the immune system of the subject may prevent infection by an antigen or a virus (e.g., an influenza virus) and thereby treat the subject or prevent the subject from disease, illness and, possibly, death.

In an additional embodiment, the invention is a composition comprising a flagellin component that is at least a portion of a flagellin, wherein the flagellin component includes at least one cysteine residue and whereby the flagellin component activates a Toll-like Receptor 5.

"Flagellin component," as used herein, means at least part of or the entirety of a flagellin.

"Activates," when referring to a Toll-like Receptor (TLR), means that the component (e.g., a flagellin component or a Toll-like Receptor agonist component) stimulates a response associated with a TLR. For example, bacterial flagellin activates TLR5 and host inflammatory responses (Smith, K.D., *et al., Nature Immunology 4*:1247-1253 (2003)). Bacterial lipopeptide activates TLR1; Pam3Cys, Pam2Cys activate TLR2; dsRNA activates TLR3; LBS (LPS-binding protein) and LPS (lipopolysaccharide) activate TLR4; imidazoquinolines (anti-viral compounds and ssRNA) activate TLR7; and bacterial DNA (CpG DNA) activates TLR9. TLR1 and TLR6 require heterodimerization with TLR2 to recognize ligands (e.g., TLR agonists, TLR antagonists). TLR1/2 are activated by triacyl lipoprotein (or a lipopeptide, such as Pam3Cys), whereas TLR6/2 are activated by diacyl lipoproteins (e g., Pam2Cys), although there may be some cross-recognition. In addition to the natural ligands, synthetic small molecules including the imidazoquinolines, with subclasses that are specific for TLR7 or TLR8 can activate both TLR7 and TLR8.

There are also synthetic analogs of LPS that activate TLR4, such as monophosphoryl lipid A [MPL].

TLR activation can result in signaling through MyD88 and NF-κB. There is some evidence that different TLRs induce different immune outcomes. For example, Hirschfeld, *et al. Infect Immun 69*:1477-1482 (2001)) and Re, *et al. J Biol Chem 276*:37692-37699 (2001) demonstrated that TLR2 and TLR4 activate different gene expression patterns in dendritic cells. Pulendran, *et al J Immunol 167*:5067-5076 (2001)) demonstrated that these divergent gene expression patterns were recapitulated at the protein level in an antigen-specific response, when lipopolysaccharides that signal through TLR2 or TLR4 were used to guide the response (TLR4 favored a Th1-like response with abundant IFNγ secretion, while TLR2 favored a Th2-line response with abundant IL-5, IL-10, and IL-13 with lower IFNγ levels). There is redundancy in the outcome of signaling through different TLRs.

Activation of TLRs can result in increased effector cell activity that can be detected, for example, by measuring IFNγ-secreting CD8+ cells (e.g., cytotoxic T-cell activity); increased antibody responses that can be detected by, for example, ELISA, virus neutralization, and flow cytometry (Schnare, M., *et al., Nat Immunol 2*:947 (2001); Alexopoulou, *L., et al., Nat Med 8*:878 (2002); Pasare, C., *et al., Science 299*:1033(2003); Napolitani, G., *et al., Nat Immunol 6*:769 (2005); and Applequist, S.E., *et al., J Immunol 175*:3882 (2005)).

The composition comprising a flagellin component that is at least a portion of a flagellin, wherein the flagellin component includes at least one cysteine residue and whereby the flagellin component activates a Toll-like Receptor 5 can further include at least a portion of at least one member selected from the group consisting of a Toll-like Receptor 1 agonist, a Toll-like Receptor 2 agonist (e.g., Pam3Cys, Pam2Cys), a Toll-like Receptor 3 agonist, a Toll-like Receptor 4 agonist, a Toll-like Receptor 6 agonist, a Toll-like Receptor 7 agonist, a Toll-like Receptor 8 agonist a Toll-like Receptor 9 agonist, a Toll-like Receptor 10 agonist, a Toll-like Receptor 11 agonist and a Toll-like Receptor 12 agonist.

The a Toll-like Receptor 1 agonist, a Toll-like Receptor 2 agonist, a Toll-like Receptor 3 agonist, a Toll-like Receptor 4 agonist, a Toll-like Receptor 6 agonist, a

Toll-like Receptor 7 agonist, a Toll-like Receptor 8 agonist a Toll-like Receptor 9 agonist a Toll-like Receptor 10 agonist, a Toll-like Receptor 11 agonist and a Toll-like Receptor 12 agonist employed in the compositions, such as compositions comprising a flagellin component that is at least a portion of a flagellin, wherein the flagellin component includes at least one cysteine residue and whereby the flagellin component activates TLR5, can further include at least one additional cysteine residue.

In one embodiment, at least one cysteine residue substitutes for at least one amino acid residue in a naturally occurring flagellin amino acid sequence of the flagellin component.

The cysteine residue that substitutes for at least one amino acid residue in a naturally occurring flagellin amino acid sequence of the flagellin component can be remote to at least one amino acid of the Toll-like Receptor 5 recognition site of the flagellin component. "Toll-like Receptor 5 recognition site," means that part of the TLR5 ligand (e.g., TLR5 agonist) that interacts with TLR5 to mediate a cellular response. "Toll-like Receptor 5 recognition site" is also referred to as a "Toll-like Receptor 5 activation site" and a "Toll-like Receptor 5 activation domain."

Likewise, "Toll-like Receptor recognition site," means that part of the Toll-like Receptor ligand (e.g., a Toll-like Receptor agonist) that interacts with its respective TLR to mediate a cellular response. "Toll-like Receptor recognition site" is also referred to as a "Toll-like Receptor activation site" and a "Toll-like Receptor activation domain."

The cysteine residue that substitutes for at least one amino acid residue in a naturally occurring flagellin amino acid sequence of the flagellin component can also be remote to at least one amino acid of the flagellin component involved in binding to the toll like receptor 5.

In another embodiment, the flagellin component includes at least a portion of a naturally occurring flagellin amino acid sequence in combination with the cysteine residue.

The cysteine residue used in combination (also referred to herein as "added cysteine") with at least a portion of a naturally occurring flagellin amino acid sequence can be at least one member selected from the group consisting of the

amino-terminal amino acid of the flagellin component or the Toll-like Receptor agonist component and the carboxy-terminal amino acid of the flagellin component or the Toll-like Receptor agonist component. The cysteine residue used in combination with at least a portion of a naturally occurring flagellin amino acid sequence can be remote to at least one amino acid of the Toll-like Receptor 5 recognition site of the flagellin component or remote to at least one amino acid of the Toll-like Receptor recognition site of the Toll-like Receptor agonist component.

The cysteine residue used in combination with at least a portion of a naturally occurring flagellin amino acid sequence can be remote to at least one amino acid of the flagellin component involved in binding to the Toll-like Receptor 5 or remote to at least one amino acid of a Toll-like Receptor agonist component involved in binding to the Toll-like Receptor.

In another embodiment, the flagellin component lacks at least a portion of a hinge region of a naturally occurring flagellin amino acid sequence.

The composition comprising a flagellin component that is at least a portion of a flagellin, wherein the flagellin component includes at least one cysteine residue and whereby the flagellin component activates a Toll-like Receptor 5 can further include at least a portion of at least one antigen (e.g., an influenza antigen, such as an influenza A, B or C antigen).

The antigen can be an essentially hydrophobic antigen, such as a maturational cleavage site antigen of HA. The maturational cleavage site antigen can be at least one member selected from the group consisting of (SEQ ID NOS: 530, 532, 533, 534, 535, 600, 601, 602, 603 and NVPEKQTRGIFGAIAGFIE (H3) (SEQ ID NO: 796), NIPSIQSRGLFGAIAGFIE (H1) (SEQ ID NO: 797), PAKLLKERGFFGAIAGFLE (FLU B) (SEQ ID NO: 798), RERRRKKRGLFGAIAGFIE (H5) (SEQ ID NO: 799), RGLXGAIAGFIE (SEQ ID NO: 821), RGLXGAIAGFIE (SEQ ID NO: 822).

"Essentially hydrophobic," as used herein, means that the antigen has limited solubility in an aqueous solution or environment. The hydrophobic nature of peptides or proteins can be determined and compared, for example, by the Kyte-Doolitle hydrophobicity scale (Kyte, J., *et al.*, *Mol. Biol.* 157: 105-132 (1982), (1982)), which assigns a numerical value for each of the 20 amino acids according to

relative hydrophobicity. A positive value indicates a hydrophobic amino acid. A negative value indicates a hydrophilic amino acid. The average of these values for an individual peptide or polypeptide (calculated by adding the individual hydrophobicity values for each amino acid of the polypeptide or protein and dividing the total value by the number of amino acids in the polypeptide or protein) provides an index of overall hydrophobicity known as the "grand average of hydropathicity," also referred to as "GRAVY." A GRAVY value greater than zero indicates the protein or peptide is essentially hydrophobic. A GRAVY value less than zero indicates the protein or peptide is essentially hydrophilic. The individual hydrophobicity values, for the 20 naturally occurring amino acids according to the Kyte-Doolittle scale (Kyte, J., *et al., Mol. Biol. 157*: 105-132 (1982)), are as follows:

Alanine	1.8
Arginine	-4.5
Asparagine	-3.5
Aspartic acid	-3.5
Cysteine	2.5
Glutamine	-3.5
Glutamic acid	-3.5
Glycine	-0.4
Histidine	-3.2
Isoleucine	4.5
Leucine	3.8
Lysine	-3.9
Methionine	1.9
Phenylalanine	2.8
Proline	-1.6
Serine	-0.8
Threonine	-0.7
Tryptophan	-0.9
Tyrosine	-1.3
Valine	4.2

The hydrophobicity of antigens, such as protein or peptide antigens, for use in the compositions and methods of the invention can be determined by calculating a GRAVY score based on the hydrophobicity values of the above Kyte-Doolittle scale. For example, GRAVY scores, which indicate essentially hydrophobic peptides, were calculated by addition of the hydrophobicity values (*supra*) of individual amino acids for the following maturational cleavage site amino peptides:

Maturational Cleavage Site	GRAVY
NVPEKQT**RGIFGAIAGFIE** (A/H3N2) (SEQ ID NO: 800)	0.053
NVPQIES**RGLFGAIAGFIE** (A/H2N1) (SEQ ID NO: 801)	0.453
NIPSIQS**RGLFGAIAGFIE** (A/H1N1) (SEQ ID NO: 802)	0.611
RGLFGAIAGFIE (Influenza A conserved region) (SEQ ID NO: 803)	1.067
PAKLLKE**RGFFGAIAGFLE** (Influenza B) (SEQ ID NO: 804)	0.400
RGFFGAIAGFLE (Influenza B conserved region) (SEQ ID NO: 805)	0.925

Likewise, GRAVY scores, which indicate essentially hydrophilic peptides, were calculated for the following peptides:

Amino Acid Sequence	GRAVY
human influenza M2e (SEQ ID NO: 806)	
SLLTEVETPIRNEWGSRSNDSSDP	-1.129
Vietnam influenza M2e (SEQ ID NO: 807)	
GSGAG SLLTEVETPTRNEWECRCSDSSDP	-0.907
Hong Kong influenza M2e (SEQ ID NO: 808)	
GSGAGSLLTEVETLTRNGWGCRCSDSSDP	-0.507
West Nile Virus E peptide 001 (SEQ ID NO: 809)	
LTSGHLKCRVKMEKLQLKGT	-0.475
Dengue 2 E peptide (SEQ ID NO: 810)	
EAEPPFGDSYIIIGVEPGQLKLNWFKK	-0.333
BCRABL wt peptide (SEQ ID NO: 811)	
SLLTEVETPIRNEWGSRSNDSSDP	-1.129

An antigen can be any molecule (e.g., protein, peptide, glycoprotein, glycopeptide, carbohydrate, lipid, lipopeptide, polysaccharide) that can be recognized by the components of the immune system regardless of whether it can trigger activation of the immune system. The antigen can be a fragment or portion of a naturally occurring antigen or a synthetic molecule that mimics the naturally occurring antigen or a portion of the naturally occurring antigen.

The antigen can be a viral antigen. A "viral antigen," as used herein, refers to any portion of a virus (e.g., influenza virus, flavivirus) that generates an immune response in a subject either when employed in combination with a TLR agonist (e.g.,

a flagellin, Pam2Cys, Pam3Cys) or in the absence of a TLR agonist. The viral antigen can be a portion or a fragment of a naturally occurring virus or a synthetic molecule that mimics a naturally occurring virus, such as a recombinant or synthetic protein (e.g., influenza virus, flavivirus), peptide, lipid, carbohydrate, that generates an immune response in the subject. The influenza antigen can include at least one member selected from the group consisting of an influenza A antigen, influenza B antigen and an influenza C antigen. The influenza antigen can be an influenza virus integral membrane protein, such as HA, or a protein portion of HA, such as HA1-1 and HA1-2.

The antigen can be at least a portion of an influenza Matrix 2 (M2) protein, including at least a portion of the ectodomain of an influenza M2 protein (M2e).

Matrix protein 2 (M2 or M2 protein) is a proton-selective integral membrane ion channel protein of the influenza A virus. M2 is abundantly expressed at the plasma membrane of virus-infected cells, but is generally underexpressed by virions. For example, a portion of an M2 sequence of influenza A is SEQ ID NO: 508, which is encoded by SEQ ID NO: 509. The native form of the M2 protein is a homotetramer (i.e., four identical disulfide-linked M2 protein molecules). Each of the units are helices stabilized by two disulfide bonds. M2 is activated by low pH. Each of the M2 protein molecules in the homotetramer consists of three domains: a 24 amino acid outer or N (amino)-terminal domain (e.g., SEQ ID NO: 510; also referred to herein as a "human consensus sequence"), which is encoded by SEQ ID NO: 511; a 19 hydrophobic amino acid transmembrane region, and a 54 amino acid inner or C (carboxy)-terminal domain. The M2 protein can vary depending upon the influenza viral subtype (e.g., H1 and H5 subtypes of influenza A) and influenza viral source (e.g., Puerto Rico, Thailand, New York, Hong Kong), as shown, for example, in exemplary amino-terminal sequences of M2 proteins (SEQ ID NOS: 544-556 and 570-578) and as described in PCT/US2005/046662 (WO2006/069262).

The M2 protein has an important role in the life cycle of the influenza A virus. It is important in the uncoating stage where it permits the entry of protons into the viral particle, which lowers the pH inside the virus, resulting in dissociation of the viral matrix protein M1 from the ribonucleoprotein RNP. As a consequence,

the virus coat is removed and the contents of the virus are released from the endosome into the cytoplasm of the host cell for infection.

The function of the M2 channel can be inhibited by antiviral drugs, such as amantadine and rimantadine, which prevent the virus from infecting the host cell. Such antiviral drugs usually bind the transmembrane region of the M2 protein and sterically block the ion channel created by the M2 protein, which prevents protons from entering and uncoating the virion.

The M2 protein for use in the compositions and methods of the invention can that include at least a portion of SEQ ID NO: 510 encoded by SEQ ID NO: 511 or at least a portion of SEQ ID NO: 544, encoded by SEQ ID NO: 603. The M2 protein can further include at least one member selected from the group consisting of SEQ ID NO: 512, SEQ ID NO: 516, SEQ ID NO: 531; SEQ ID NO: 536 (Flu A H5N1 M2e, 2004 Viet Nam Isolate with serine replacing cysteine); SEQ ID NO: 537 (Flu A H5N1 M2e, 2004 Viet Nam Isolate); SEQ ID NO: 538 (Flu A H5N1 M2e, Hong Kong 97 Isolate with serine replacing cysteine); SEQ ID NO: 539 (Flu A H5N1 M2e, Hong Kong 97 Isolate); SEQ ID NO: 540 (Flu A H7N2 M2e Chicken/New York 95 Isolate with serine replacing cysteine); SEQ ID NO: 541 (Flu A H7N2 M2e, Chicken/ New York 95 Isolate); SEQ ID NO: 542 (Flu A H9N2 M2e, Hong Kong 99 Isolate with serine replacing cysteine); and SEQ ID NO: 543 (Flu A, Hong Kong 99 Isolate). Certain cysteine residues, for example, amino acids 16 and 18 of SEQ ID NO: 537; amino acids 17 and 19 of SEQ ID NOS: 539, 541 and 543 in the naturally occurring sequence of at least a portion of M2 protein are replaced with a serine (see, SEQ ID NOS: 538, 540, 542 and 544, respectively).

The compositions comprising a flagellin component that is at least a portion of a flagellin, wherein the flagellin component includes at least one cysteine residue and whereby the flagellin component activates a Toll-like Receptor 5 that further includes an antigen (e.g., a maturational cleavage site peptide, protein portion of HA, such as HA1-1, HA1-2) can be co-administered or admixed with another flagellin component that is at least a portion of a flagellin, wherein the flagellin component includes at least one cysteine residue and whereby the flagellin component activates a Toll-like Receptor 5 that further includes a different, distinct antigen (e.g., M2e) or a fusion protein of a TLR agonist and an antigen (e.g.,

STF2.HA1-1, STF2.HA1-2, STF2Δ.HA1-1, STF2Δ.HA1-2, STF2.M2e, STF2.4xM2e, STF2Δ.M2e, STF2Δ.4xM2e).

The antigen included in the compositions and employed in the methods of the invention can be at least a portion of a microbial-related antigen, for example, antigens of pathogens, such as viruses, fungi or bacteria. Antigens can also be microorganism-related, and other disease-related antigens, such as antigens associated with allergies and cancer. The antigen can also be at least a portion of an antigen of a bacteria, a virus, a fungi, a yeast, a protozoa, a metazoa, a tumor, a malignant cell, a plant cell, an animal cell, a hormone and an amyloid-ß peptide. The antigen can be a peptide, a polypeptide, a lipoprotein, a glycoprotein and a mucoprotein.

The antigen included in the compositions and employed in the methods of the invention can be a pathogen-related antigen. "Pathogen-related antigen," as used herein, refers to any molecule (e.g., protein, peptide, carbohydrate, lipoprotein, polysaccharide) that is associated with a pathogen. Exemplary pathogen-related antigens include, for example, antigens of vaccinia, avipox virus, turkey influenza virus, bovine leukemia virus, feline leukemia virus, avian influenza, chicken pneumovirosis virus, canine parvovirus, equine influenza, FHV, Newcastle Disease Virus (NDV), Chicken/Pennsylvania/1/83 influenza virus, infectious bronchitis virus; Dengue virus, measles virus, Rubella virus, pseudorabies, Epstein-Barr Virus, HIV, SIV, EHV, BHV, HCMV, Hantaan, C. *tetani*, mumps, Morbillivirus, Herpes Simplex Virus type 1, Herpes Simplex Virus type 2, Human cytomegalovirus, Hepatitis A Virus, Hepatitis B Virus, Hepatitis C Virus, Hepatitis E Virus, Respiratory Syncytial Virus, Human Papilloma Virus, Influenza Virus, *Salmonella*, *Neisseria*, *Borrelia*, *Chlamydia*, *Bordetella*, *Plasmodium* (e.g., *Plasmodium falciparum* and *Plasmodium vivax*), *Toxoplasma*, *Cryptococcus*, *Streptococcus*, *Staphylococcus*, *Haemophilus*, *Diptheria*, *Tetanus*, *Pertussis*, *Escherichia*, *Candida*, *Aspergillus*, *Entamoeba*, *Giardia* and *Trypanasoma*.

The antigen included in the compositions and employed in the methods of the invention can include bacterial capsular antigens. Exemplary bacterial capsular antigens include, for example, at least one member selected from the group consisting of a Group B *Streptococcus*, including *Streptococcus agalactiae*, capsular

polysaccharides type Ia, Ib, Ia/c, II, III, IV, V and VI; *Streptococcus pneumoniae* capsular polysaccharides types 1, 2, 3, 4, 5, 6A, 6B, 7F, 9N, 9V, 10A, 11A, 12F, 14, 15B, 17F, 18C, 19A, 20, 22F, 23F, and 33F or any combination of these (typing according to the Danish system); *Stapylococcus aureus* capsular polysaccharides type 5, 8 and 336; *Bacillus anthracis* vegetative state capsular poly gamma D glutamic acid; *Mycobacteria* tuberculosis lipo-arabino mannan capsule; *Plasmodium flaciparum* surface oligosaccharide; *Neisseria meningitidis* capsular polysaccharides of group A, B, C, X, Y and W135 or any combination thereof.

Compositions comprising the flagellin component that is at least a portion of a flagellin, wherein the flagellin component includes at least one cysteine residue and whereby the flagellin component activates a Toll-like Receptor 5 can include an antigen that is produced by chemical modification of the side-chains of the antigen (e.g., capsular antigens, pathogen-related antigens, influenza antigens, flavivirus antigens) in a manner that generates an active group that is coupled (also referred to herein as chemical conjugation) to a compatible reactive group created on the substituted flagellins of the invention (see *infra*). Derivatization (modification) of the capsular material is carried out so that a small minority of the side-chains are modified (about 5%) to avoid destruction of the majority of immunoreactive structures on the capsular material.

The antigen included in the compositions and employed in the methods of the invention can be at least a portion of at least one member selected from the group consisting of a West Nile viral protein, a Langat viral protein, a Kunjin viral protein, a Murray Valley encephalitis viral protein, a Japanese encephalitis viral protein, a Tick-borne encephalitis viral protein, Dengue 1 viral protein, Dengue 2 viral protein, Dengue 3 viral protein, Dengue 4 viral protein, hepatitis C viral protein and a Yellow fever viral protein (see, for example, PCT/US2006/001623 (WO2006/078657)).

The genus flavivirus is in the virus family *Flaviviridae* and consists of about 70 viruses. Mosquito or ticks transmit most of these viruses. Several flaviviruses are significant human pathogens, including the four dengue viruses (Den1, Den2, Den3 and Den4), yellow fever (YF), Japanese encephalitis (JE), West Nile (WN, also referred to herein as "WNV") and Tick-borne encephalitis (TBE) (Weaver S.C.,

et al., Nat Rev Microbiol 10: 789-801 (2004)). The flavivirus-genus is divided into a number of serogroups based on cross-neutralization tests, including the dengue serogroup that contains four serologically and genetically distinct viruses termed DEN-1, DEN-2, DEN-3 and DEN-4.

Flaviviruses are small, enveloped viruses with icosahedral capsids. The flavivirus genome is a single-stranded positive-sense RNA (about 11 kb) that is directly translated by the host cell machinery following infection. The viral genome is translated as a single polypeptide that undergoes co- and post-translational cleavage by viral and cellular enzymes to generate three structural proteins of the flavivirus (the capsid (C), the membrane (M) and the envelope (E) proteins); and seven nonstructural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5) (Weaver, et al., *Annu Rev Microbiol 1990*:44-649 (2004)). The viral capsid is composed of the C-protein, while both the M- and envelope proteins are located on the envelope surface of the virion (Weaver, S.C., *et al., Nat. Rev. Microbiol. 10*:789-801 (2004); Chambers *et al., Annu Rev. Microbiol. 44*: 649-688 (1990)). A major immunogen for flaviviruses is the membrane envelope protein.

A flavivirus can enter a host cell when the viral envelope protein binds to a receptor and responds by conformational rearrangement to the reduced pH of an endosome. The conformational change induces fusion of viral and host-cell membranes.

The envelope of a flavivirus may function as a receptor binding protein and to facilitate fusion of the virus and host cell membrane. Envelope proteins of flaviviruses have common structural (domains I, II and III) and functional features (receptor binding of virus and host cell and fusion functions) and are class II fusion glycoproteins (Lescar *et al., Cell 105*:137-148 (2001)).

In the pre-fusion conformation, envelope proteins form homodimers on the outer surface of the virus particles (Rey, *et al., Nature 375*:291-298); Kuhn, *et al., Cell 108*:717-725 (2002); Mukhopadhyay, *et al., Science 302*:248 (2003)). Each envelope protein monomer folds into three structural domains (domains I, II and III) predominantly composed of β-strands. Domain I (also referred to herein as "I" or "DI") is centrally located in the structure and has an N-glycosylation site in glycosylated envelope proteins. Domain II (also referred to herein as "II" or "DII")

of the envelope protein promotes dimerization and has a fusion loop that inserts into the target host membrane during the pH-dependent fusion of the virus (Modis, *et al.*, *Nature 427:*313-319 (2004); Bressanelli, *et al.*, *EMBO J 23:*728-738 (2004)). Domain III (also referred to herein as "III" or "DIII") is at the carboxy-terminus of the envelope protein. Domain III is also referred to as "domain B" in earlier antigenic mapping studies. Domain III has several epitopes that can elicit virus-neutralizing antibodies (Roehrig, *Adv Virus Res 59:*141-175 (2003)).

Domain I of the Tick-borne encephalitis envelope protein corresponds to amino acids 1-51, 137-189 and 285-302 of SEQ ID NO: 777; domain II of the Tick-borne encephalitis envelope protein of SEQ ID NO: 777 corresponds to amino acids 52-136 and 190-284; and domain III corresponds to amino acids 303-395 of SEQ ID NO: 777. (Rey, F.A., *et al.*, *Nature 375*:291-298 (1995)). SEQ ID NO: 777 is encoded by SEQ ID NO: 778. Domain I of the Dengue 2 flavivirus envelope protein corresponds to amino acids 1-52, 132-193 and 280-296 of SEQ ID NO: 763; domain II corresponds to amino acids 53-131 and 194-279 of SEQ ID NO: 763; and domain III corresponds to amino acids 297-495 of SEQ ID NO: 763 (Modis, Y., *et al.*, *Nature 427*:313-319 (2004)). The location of domains I, II and III of other flavivirus (e.g., West Nile virus, Japanese encephalitis, Dengue 1 virus, Dengue 3 virus and Dengue 4 virus) is based on homology of the Tick-borne encephalitis envelope protein domains and the Dengue 2 envelope protein domains. Thus, reference herein to domains of flavivirus proteins, in particular, flaviviruses other than Tick-borne encephalitis flavivirus envelope proteins and Dengue 2 flavivirus envelope proteins, are based on homology to domains in the Tick-borne encephalitis flavivirus envelope protein and the Dengue 2 flavivirus envelope protein.

The domain III of the envelope protein of the DEN flavivirus encodes the majority of the flavivirus type-specific contiguous critical/dominant neutralizing epitopes (Roehring, J.T., *Adv. Virus Res. 59*:141 (2003)), including the four DEN (DEN1, DEN2, DEN3, DEN4) viruses. Flavivirus envelope proteins are highly homologous. Exemplary envelope protein sequences are SEQ ID NOS: 642, 763, 765, 767, 769 and 774.

West Nile virus (WNV) is a single-stranded positive sense RNA envelope virus. It was first isolated and identified in the West Nile region of Uganda in 1937 from a febrile female adult (Smithburn, *et al.*, *Am J Trop Med Hyg 3*:9-18 (1954)).

Japanese encephalitis (JE) virus is localized in Asia and northern Australia (about 50,000 cases with about10,000 deaths annually).

The Dengue (DEN) disease is caused by four mosquito-borne, serologically related flaviviruses known as DEN-1 (also referred to herein as "Den1" or Den 1"), DEN-2 (also referred to herein as "Den2" or "Den 2"), DEN-3 (also referred to herein as "Den3" or "Den 3"), and DEN-4 (also referred to herein as "Den4" or Den 4"). The compositions, fusion proteins and polypeptides of the invention can include Den 1 SEQ ID NO.: 623; Den 1 PR 94 (Puerto Rico, 1994) SEQ ID NO: 624; Den 3 SEQ ID NO: 626; and Den 4 SEQ ID NO: 627. SEQ ID NOS: 623, 624, 625, 626 and 627 are portions of domain III of Den1, Den2, Den3 and Den4 flaviviruses.

"EI," "EII," and "EIII," as used herein, refer to domains I, II and III, respectively, of the West Nile flavivirus envelope protein. "JEI," "JEII," and "JEIII," as used herein, refer to domains I, II and III, respectively, of the Japanese encephalitis flavivirus envelope protein. "Den1 I," "Den1 II," and "Den1 III," as used herein refer to domains I, II and III, respectively, of the Dengue 1 flavivirus envelope protein. Likewise, designations for the domains of envelope proteins of other flaviviruses are referenced by the flavivirus name followed by the domain number (e.g., (Tick-borne) TBI (Tick-borne), TBII, TBIII, Den2 I, Den2 II, Den2 III).

The portion of an envelope protein of a flavivirus can include at least one member selected from the group consisting of at least a portion of domain I, at least a portion of domain II and at least a portion of domain III. When a domain is designated with a "+," for example "EIII+" or "JEIII+," the portion of the envelope protein referenced as "III" is one component of the total of that domain plus at least one of at least a portion of either or both of domains I and II. For example, "EIII+," as used herein, means the compositions, fusion proteins and polypeptides of the invention include domain III and at least a portion of domain I. "EIII+" is also referred to as "EI/III." "JEIII+" is also referred to as "JEI/III." Similarly, when

compositions include domains of envelope proteins of flavivirus, the domains can be any combination of domains I, II, and III and can be designated based on the domain. For example, EI/II includes domain I and II of the West Nile flavivirus. The absence of a "+" in reference to a domain (e.g., EIII, JEIII, Den1 III) of an envelope protein employed in the compositions, fusion proteins and polypeptides of the invention means that the composition, fusion protein and polypeptide includes the referenced domain. For example, "Den1 III" means the compositions, fusion proteins and compositions include domain III, not domains I and II, of the Dengue 1 virus.

The West Nile viral envelope protein can include at least a portion of at least one member selected from the group consisting of SEQ ID NO: 610, which is an EIII+ amino acid sequence, the italicized amino acids are domain I of the envelope protein and the remaining sequence is domain III of the envelope protein; SEQ ID NO: 611, West Nile virus, Stanford, CT, also referred to as "West Nile S"; SEQ ID NO: 612, West Nile virus, New York, NY, also referred to as "West Nile NY"; and SEQ ID NO: 613, SEQ ID NO: 610 is encoded by SEQ ID NO: 614.

The Langat virus envelope protein for use in the compositions, fusion proteins and polypeptides of the invention can include at least a portion of SEQ ID NO: 615. The Kunjin virus envelope protein can include at least a portion of SEQ ID NO: 616. The Murray Valley encephalitis envelope protein can include at least a portion of SEQ ID NO: 617. The Japanese encephalitis envelope protein can include at least one member selected from the group consisting of at least a portion of SEQ ID NO: 618 and SEQ ID NO: 619. The Tick-borne encephalitis envelope protein can include at least a portion of SEQ ID NO: 620. The Yellow fever virus envelope protein can include at least a portion of SEQ ID NO: 621. The envelope protein of a flavivirus can include at least a portion of at least one member selected from the group consisting of SEQ ID NO: 622 and SEQ ID NO: 643. SEQ ID NOS: 615, 616, 617, 618, 619, 620, 621, 622 and 643 are portions of domain III of the viral envelope protein.

The antigen can be chemically conjugated to flagellin components and Toll-like Receptor agonist components. Chemical conjugation (also referred to herein as "chemical coupling") can include conjugation by a reactive group, such as a thiol

group (e.g., a cysteine residue) or by derivatization of a primary (e.g., a amino-terminal) or secondary (e.g., lysine) group. Different crosslinkers can be used to chemically conjugate TLR ligands (e.g., TLR agonists) to proteins (e.g., antigens, compositions of the invention, HA and M2e constructs of the invention) or other molecules (e.g., nucleic acids, polysaccharides). Exemplary cross linking agents are commerically available, for example, from Pierce (Rockland, Ill). Methods to chemically conjugate the antigen to the flagellin component are well-known and include the use of commercially available cross-linkers, such as those described herein.

For example, conjugation of peptide or protein antigens to a flagellin component or a Toll-like Receptor agonist component of the invention can be through at least one cysteine residue of the flagellin component or the Toll-like Receptor component and at least one cysteine residue of a protein (e.g., an influenza antigen, such as HA, M2e) employing established techniques. The protein can be derivatized with a homobifunctional, sulfhydryl-specific crosslinker; the protein is then desalted to remove the unreacted crosslinker; and then the peptide or protein partner added and conjugated via at least one cysteine residue cysteine. Exemplary reagents for use in the conjugation methods can be purchased commercially from Pierce (Rockland, Ill), for example, BMB (Catalog No: 22331), BMDB (Catalog No: 22332), BMH (Catalog No: 22330), BMOE (Catalog No: 22323), $BM[PEO]_3$ (Catalog No: 22336), $BM[PEO]_4$ (Catalog No:22337), DPDPB (Catalog No: 21702), DTME (Catalog No: 22335), HBVS (Catalog No: 22334).

Alternatively, cysteine-containing proteins and antigens can also be conjugated to lysine residues on flagellin components, flagellin, Toll-like Receptor agonist components and Toll-like Receptor agonists of the invention. A protein containing no cysteine residues is derivatized with a heterobifunctional amine and sulfhydryl-specific crosslinker. After desalting, the cysteine-containing partner is added and conjugated. Exemplary reagents for use in the conjugation methods can be purchased from Pierce (Rockland, Ill), for example, AMAS (Catalog No: 22295), BMPA (Catalog No. 22296), BMPS (Catalog No: 22298), EMCA (Catalog No: 22306), EMCS (Catalog No: 22308), GMBS (Catalog No: 22309), KMUA (Catalog No: 22211), LC-SMCC (Catalog No: 22362), LC-SPDP (Catalog No: 21651), MBS

(Catalog No: 22311), SATA (Catalog No: 26102), SATP (Catalog No: 26100), SBAP (Catalog No: 22339), SIA (Catalog No: 22349), SIAB (Catalog No: 22329), SMCC (Catalog No: 22360), SMPB (Catalog No: 22416), SMPH (Catalog No. 22363), SMPT (Catalog No: 21558), SPDP (Catalog No: 21857), Sulfo-EMCS (Catalog No: 22307), Sulfo-GMBS (Catalog No: 22324), Sulfo-KMUS (Catalog No: 21111), Sulfo-LC-SPDP (Catalog No: 21650), Sulfo-MBS (Catalog No: 22312), Sulfo-SIAB(Catalog No: 22327), Sulfo-SMCC (Catalog No: 22322), Sulfo-SMPB (Catalog No: 22317), Sulfo-LC-SMPT (Catalog No.: 21568).

Additionally, or alternatively, peptide or protein antigens can also be conjugated to flagellin components or Toll-like Receptor agonist components of the invention via at least one lysine residue on both conjugate partners. The two conjugate partners are combined along with a homo-bifunctional amine-specific crosslinker. The appropriate hetero-conjugate is then purified away from unwanted aggregates and homo-conjugates. Exemplary reagents for use in the conjugation methods can be purchased from Pierce (Rockland, Ill), for example, BSOCOES (Catalog No: 21600), BS_3 (Catalog No: 21580), DFDNB (Catalog No: 21525), DMA (Catalog No: 20663), DMP (Catalog No: 21666), DMS (Catalog No: 20700), DSG (Catalog No: 20593), DSP (Catalog No: 22585), DSS (Catalog No: 21555), DST (Catalog No: 20589), DTBP (Catalog No: 20665), DTSSP (Catalog No: 21578), EGS (Catalog No: 21565), MSA (Catalog No: 22605), Sulfo-DST (Catalog No: 20591), Sulfo-EGS (Catalog No: 21566), THPP (Catalog No: 22607).

Similarly, peptide or protein antigens can be conjugated to flagellin components or Toll-like Receptor agonist components of the invention via at least one carboxyl group (e.g., glutamic acid, aspartic acid, or the carboxy-terminus of the peptide or protein) on one partner and amines on the other partner. The two conjugation partners are mixed together along with the appropriate heterobifunctional crosslinking reagent. The appropriate hetero-conjugate is then purified away from unwanted aggregates and homo-conjugates. Exemplary reagents for use in the conjugation methods can be purchased from Pierce (Rockland, Ill), for example, AEDP (Catalog No: 22101), EDC (Catalog No: 22980) and TFCS (Catalog No: 22299).

In addition, carbohydrate antigens can be conjugated to proteins, flagellin components, flagellin, Toll-like Receptor agonist components and Toll-like Receptor agonists of the invention via at least one cysteine residue in the protein employing well-established techniques. For example, the protein is initially derivatized with a heterobifunctional crosslinker containing at least one sulfhydryl-specific group and at least one hydrazide group. The polysaccharide or oligosaccharide is then treated with an oxidizing agent such as sodium meta-periodate to generate terminal aldehyde groups. The oxidized carbohydrate is then added to the derivatized protein and conjugation to the aldehyde occurs via the hydrazide on the crosslinker. Exemplary reagents for use in the conjugation methods can be purchased from Pierce (Rockland, Ill), for example, BMPH (Catalog No: 22297), EMCH (Catalog No: 22106), KMUH (Catalog No: 22111) and PDPH (Catalog No: 22301).

Further lipopeptides can be conjugated to protein antigens, flagellin components or Toll-like Receptor components or Toll-like Receptor agonists of the invention via amino-acid side chains on the peptide chain. Similar strategies and reagents, as described above for peptide conjugation to proteins, would be employed.

TLRs can be activated by nucleic acids. For example, TLR3 is activated by double-stranded (ds) RNA; TLR7 and TLR8 are activated by single-stranded (ss) RNA; and TLR9 is activated by CpG DNA sequences. Several different techniques can be employed to conjugate Toll-like Receptor agonist components of nucleic acid TLRs to protein antigens. For example, for un-modified DNA molecule, the 5' phosphate group can be modified with the water-soluble carbodiimide EDC (Pierce; Rockford, IL, Catalog No: 22980) followed by imidazole to form a terminal phosphoylimidazolide. A terminal amine can then be substituted for this reactive group by the addition of ethylenediamine. The amine-modified nucleic acid can then be conjugated to a cysteine on a protein using a heterobifunctional maleimide (cysteine-specific) and NHS-ester (lysine-specific) crosslinker, as described above for peptide-protein conjugation.

Alternatively, or additionally, a sulfhydryl group can be incorporated at the 5' end by substituting cystamine for ethylenediamine in the second step. After

reduction of the disulfide bond, the free sulfhydryl can be conjugated to cysteines on proteins, flagellin components, flagellin, Toll-like Receptor agonist components and Toll-like Receptor agonists employing a homo-bifunctional maleimide-based crosslinker, or to lysines using a heterobifunctional crosslinker, as described above. Alternately, or additionally, synthetic oligonucleotides can be synthesized with modified bases on either the 5' or 3' end. These modified bases can include primary amine or sulfhydryl groups which can be conjugated to proteins using the methods described above. The 3' end of RNA molecules may be chemically modified to allow coupling with proteins or other macromolecules. The diol on the 3'-ribose residue may be oxidized using sodium meta-periodate to produce an aldehyde group. The aldehyde may then be conjugated to proteins using a hydrazide-containing crosslinker, such as MPBH (Pierce; Rockland, IL, Catalog No: 22305), which covalently modifies the carbonyl group, and then conjugates to free thiols on proteins via a maleimide group. Alternatively, the 3' hydroxyl group may be derivatized directly with an isocyanate-containing crosslinker, such as PMPI (Pierce; Rockland, IL, Catalog No: 28100), which also contains a maleimide group for conjugation to protein sulfhydryls.

Synthetic small-molecule TLR ligands (e.g., Toll-like Receptor agonists, Toll-like Receptor antagonists) have been identified. For example, imiquimod, which potently activates TLR7, and resiquimod, an activator of both TLR7 and TLR8. Analogs of these compounds with varying levels of potency and specificity have been synthesized. The ability to conjugate a small molecule TLR agonist to an antigen of interest can depend on the chemical nature of the TLR ligand. Some TLR ligands may have active groups that can be exploited for chemical conjugation. For example, imiquimod, as well as its analog gardiquimod (InvivoGen; San Diego, CA) and the TLR7-activating adenine analog CL087 (InvivoGen, San Diego, CA), have primary amine groups (-NH_2), which can be targets for derivatization by crosslinkers containing imidoesters or NHS esters. In another strategy, gardiquimod contains an exposed hydroxyl (-OH) group, which can be derivatized by an isocyanate-containing crosslinker, such as PMPI (Pierce; Rockland, ILL Catalog No: 28100) for subsequent crosslinking to protein sulfhydryl groups.

In addition, custom synthesis of derivatives of small-molecule TLR ligands can be arranged in order to attach novel functional groups at different postions in the molecule to facilitate crosslinking. Custom derivatives can also include groups, such as maleimides, which can then be used for direct linking to protein antigens.

Chemical conjugation of an antigen to the flagellin component can result in increased aqueous solubility of the antigen (e.g., an essentially hydrophobic antigen, such as a maturational cleavage site antigen) as a component of the composition.

The composition comprising a flagellin component that is at least a portion of a flagellin, wherein the flagellin component includes at least one cysteine residue and whereby the flagellin component activates a Toll-like Receptor 5 can include a cysteine residue in the hypervariable region of the flagellin component.

At least one cysteine residue substitutes for at least one amino acid in a naturally occurring flagellin amino acid sequence flagellin component. The cysteine residue can substitute for at least one amino acid selected from the group consisting of amino acid 1, 237, 238, 239, 240, 241 and 495 of SEQ ID NO: 812; at least one amino acid selected from the group consisting of amino acid 1, 240, 241, 242, 243, 244 and 505 of SEQ ID NO: 498; at least one amino acid selected from the group consisting of amino acid 1, 237, 238, 239, 240, 241 and 504 of SEQ ID NO: 504; at least one amino acid selected from the group consisting of amino acid 1, 211, 212, 213 and 393 of SEQ ID NO: 815; at least one amino acid selected from the group consisting of amino acid 1, 151, 152, 153, 154 and 287 of SEQ ID NO: 820; at least one amino acid selected from the group consisting of amino acid 1, 238, 239, 240, 241, 242, 243 and 497 of SEQ ID NO: 502; at least one amino acid selected from the group consisting of amino acid 1, 237, 238, 239, 240, 241 and 495 of SEQ ID NO: 812.

The flagellin component or Toll-like Receptor agonist component can include at least a portion of a naturally occurring flagellin amino acid sequence in combination with the cysteine residue.

The composition of the invention wherein the cysteine residue substitutes for at least one amino acid in a naturally occurring flagellin amino acid sequence flagellin component, or wherein the flagellin component includes at least a portion of a naturally occurring flagellin amino acid sequence in combination with the

cysteine residue, can activate a Toll-like Receptor 5. For example, a cysteine residue can be placed within the D1/D2 domain proximate to the amino-terminus and carboxy-terminus, remote to the TLR5 recognition site (see, for example, Figures 75 and 76). Alternatively, or additionally, the cysteine residue can be placed at the distal point of the hypervariable domain (see, for example, Figures 75 and 76) at about amino acid 237, about 238, about 239, about 240 and about 241 of SEQ ID NO: 812. Substituting polar or charged amino acids is preferable to substituting hydrophobic amino acids with cysteine residues. Substitution within the TLR5 recognition site is least preferable.

Flagellin from *Salmonella typhimurium* STF1 (FliC) is depicted in SEQ ID NO: 812 (Accession No: P06179). The TLR5 recognition site is amino acid about 79 to about 117 and about 408 to about 439. Cysteine residues can substitute for or be included in combination with amino acid about 408 to about 439 of SEQ ID NO: 812; amino acids about 1 and about 495 of SEQ ID NO: 812; amino acids about 237 to about 241 of SEQ ID NO: 812; and/or amino acids about 79 to about 117 and about 408 to about 439 of SEQ ID NO: 812.

Salmonella typhimurium flagellin STF2 (FljB) is depicted in SEQ ID NO: 498. The TLR5 recognition site is amino acids about 80 to about 118 and about 420 to about 451 of SEQ ID NO: 498. Cysteine residues can substitute for or be included in combination with amino acids about 1 and about 505 of SEQ ID NO: 498; amino acids about 240 to about 244 of SEQ ID NO: 498; amino acids about 79 to about 117 and/or about 419 to about 450 of SEQ ID NO: 498.

Salmonella muenchen flagellin is depicted in SEQ ID NO: 504 (Accession No: #P06179). The TLR5 recognition site is amino acids about 79 to about 117 and about 418 to about 449 of SEQ ID NO: 504. Cysteine residues can substitute for or be included in combination with amino acids about 1 and about 504 of SEQ ID NO: 504; about 237 to about 241 of SEQ ID NO: 504; about 79 to about 117; and/or about 418 to about 449 of SEQ ID NO: 504.

Escherichia coli flagellin is depicted in SEQ ID NO: 502 (Accession No: P04949). The TLR5 recognition site is amino acids about 79 to about 117 and about 410 to about 441 of SEQ ID NO: 502. Cysteine residues can substitute for or be included in combination with amino acids about 1 and about 497 of SEQ ID NO:

502; about 238 to about 243 of SEQ ID NO: 502; about 79 to about 117; and/or about 410 to about 441 of SEQ ID NO: 502.

Pseudomonas auruginosa flagellin is depicted in SEQ ID NO: 815. The TLR5 recognition site is amino acids about 79 to about 114 and about 308 to about 338 of SEQ ID NO: 815. Cysteine residues can substitute for or be included in combination with amino acids about 1 and about 393 of SEQ ID NO: 815; about 211 to about 213 of SEQ ID NO: 815; about 79 to about114; and/or about 308 to about 338 of SEQ ID NO: 815.

Listeria monocytogenes flagellin is depicted in SEQ ID NO: 820. The TLR5 recognition site is amino acids about 78 to about 116 and about 200 to about 231 of SEQ ID NO: 820. Cysteine residues can substitute for or be included in combination with amino acids about 1 and about 287 of SEQ ID NO: 820; about 151 to about 154 of SEQ ID NO: 820; about 78 to about 116; and/or about 200 to about 231 of SEQ ID NO: 820.

Experimentally defined TLR5 recognition sites on STF2 have been described (see, for example, Smith, K.D., *et al., Nature Immunology 4*:1247-1253 (2003) at amino acids about 79 to about 117 and about 420 to about 451. In addition, Smith, K.D., *et al., Nature Immunology 4*:1247-1253 (2003), based on sequence homology, identified TLR5 recognition sites on other flagellins, such as STF1 at amino acids about 79 to about 117, about 408 to about 439; *P. aeruginosa* at amino acids about 79 to about 117, about 308 to about 339; *L. pneumophila* at amino acids about 79 to about 117, about 381 to about 419; *E. coli* at amino acids about 79 to about 117, about 477, about 508; *S. marcesens* at amino acids about 79 to about 117, about 265-about 296; *B. subtilus* at amino acids about 77 to about 117, about 218 to about 249; and *L. monocytogenes* at amino acids about 77 to about 115, about 200 to about 231.

The high-resolution structure STF1 (FliC) (SEQ ID NO: 812) has been determined and can be a basis for analysis of TLR5 recognition by a flagellin and the location of cysteine substitutions/additions. Flagellin resembles a "boomerang," with the amino- and carboxy- termini at the end of one arm (see, for example Figure 77). The TLR5 recognition site is located roughly on the outer side of the boomerang just below the bend toward the amino- and carboxy-termini of the flagellin. The hinge region, which is not required for TLR5 recognition, is located

above the bend. The region of greatest sequence homology of flagellins is in the TLR5 recognition site. The next region of sequence homology is in the D1 and D2 domains, which include the TLR5 recognition site and the amino- and carboxy-termini. The D1 and D2 domains, with or without a linker, is STF2Δ (SEQ ID NO: 500), which can activate TLR5. The region of least sequence homology between flagellins is the hypervariable region.

It is believed that the ability of the flagellin component or Toll-like Receptor agonist component to activate TLR5 can be accomplished by maintaining the conjugation sites (cysteine residues substituted for at least one amino acid in a naturally occurring flagellin amino acid sequence flagellin component or at least a portion of a naturally occurring flagellin amino acid sequence in combination with the cysteine residue) remote from the TLR5 or TLR recognition site. For example, for STF1 (SEQ ID NO: 812), for which a high resolution structural determination is available, this may be achieved in the D1 domain, D2 domain or in the hinge region. In the D1/D2 domain the amino- and carboxy-termini can be remote (also referred to herein as "distal") from the TLR5 recognition site, and moving away from either the amino or carboxy terminus may bring the conjugation site closer to the recognition site and may interfere with TLR5 activity. In the hinge region amino acids about 237 to about 241 of SEQ ID NO: 812, are approximately at the other tip of the "boomerang" and are about the same distance from the TLR5 recognition site as the amino-and carboxy-termini. This site may also be a location that maintains TLR5 recognition.

Amino acid identity can be taken into consideration for the location of conjugation sites. Polar and charged amino acids (e.g., serine, aspartic acid, lysine) are more likely to be surface exposed and amenable to attachment of an antigen. Hydrophobic amino acids (e.g., valine, phenalalanine) are more likely to be buried and participate in structural interactions and should be avoided.

Compositions that include flagellin components with cysteine residues or Toll-like Receptor agonist components with cysteine residues activate TLR5 and can be chemically conjugated to antigens.

The compositions and methods of employing the compositions of the invention can further include a carrier protein. The carrier protein can be at least

one member selected from the group consisting of a tetanus toxoid, a *Vibrio cholerae* toxoid, a diphtheria toxoid, a cross-reactive mutant of diphtheria toxoid, a *E. coli* B subunit of a heat labile enterotoxin, a tobacco mosaic virus coat protein, a rabies virus envelope protein, a rabies virus envelope glycoprotein, a thyroglobulin, a heat shock protein 60, a keyhole limpet hemocyanin and an early secreted antigen tuberculosis-6.

The composition comprising a flagellin component that is at least a portion of a flagellin, wherein the flagellin component includes at least one cysteine residue and whereby the flagellin component activates a Toll-like Receptor 5 can include at least one lysine of the flagellin component that has been substituted with at least one member selected from the group consisting of an arginine residue, a serine residue and a histidine residue.

In an additional embodiment, the invention is a composition comprising a Toll-like Receptor agonist component that is at least a portion of a Toll-like Receptor agonist, wherein the Toll-like Receptor agonist component includes at least one cysteine residue in a position where a cysteine residue does not occur in the native Toll-like Receptor agonist, whereby the Toll-like Receptor agonist component activates a Toll-like Receptor. "Component," as used herein in reference to a Toll-like Receptor agonist component, means at least part of or the entirety of a Toll-like Receptor agonist.

In one embodiment, the cysteine residue in the composition comprising a Toll-like Receptor agonist component that is at least a portion of a Toll-like Receptor agonist, wherein the Toll-like Receptor agonist component includes at least one cysteine residue in a position where a cysteine residue does not occur in the native Toll-like Receptor agonist, whereby the Toll-like Receptor agonist component activates a Toll-like Receptor, substitutes for at least one amino acid in a naturally occurring amino acid sequence of a Toll-like Receptor agonist component. The cysteine can substitute for at least one amino acid remote to the Toll-like Receptor recognition site of the Toll-like Receptor agonist component.

In another embodiment, the Toll-like Receptor agonist component includes at least a portion of a naturally occurring Toll-like Receptor agonist amino acid sequence in combination with a cysteine residue. The cysteine residue in

combination with the naturally occurring Toll-like Receptor agonist can be remote to the Toll-like Receptor recognition site of the Toll-like Receptor agonist component.

The composition comprising a Toll-like Receptor agonist component that is at least a portion of a Toll-like Receptor agonist, wherein the Toll-like Receptor agonist component includes at least one cysteine residue in a position where a cysteine residue does not occur in the native Toll-like Receptor agonist, whereby the Toll-like Receptor agonist component activates a Toll-like Receptor can further include at least a portion of at least one antigen (e.g., an influenza antigen, such as a influenza integral membrane protein antigen, HA, HA1-1, HA1-2, M2, M2e).

In still another embodiment, the invention is a composition comprising a flagellin component that is at least a portion of a flagellin, wherein at least one lysine of the flagellin component has been substituted with at least one arginine, whereby the flagellin component activates a Toll-like Receptor 5.

"Substituted," as used herein in reference to the flagellin, flagellin component, Toll-like Receptor agonist or Toll-like Receptor agonist component, means that at least one amino acid, such as a lysine of the flagellin component, has been modified to another amino acid residue, for example, a conservative substitution (e.g., arginine, serine, histidine) to thereby form a substituted flagellin component or substituted Toll-like Receptor agonist component. The substituted flagellin component or substituted Toll-like Receptor agonist component can be made by generating recombinant constructs that encode flagellin with the substitutions, by chemical means, by the generation of proteins or peptides of at least a portion of the flagellin by protein synthesis techniques, or any combination thereof.

The lysine residue that is substituted with an amino acid (e.g., arginine, serine, histidine) can be at least one lysine residue selected from the group consisting of lysine 19, 41, 58, 135, 160, 177, 179, 203, 215, 221, 228, 232, 241, 251, 279, 292, 308, 317, 326, 338, 348, 357, 362, 369, 378, 384, 391 and 410 of SEQ ID NO: 812.

The flagellin can be a *S. typhimurium* flagellin that includes SEQ ID NO: 816. The lysine residue that is substituted with an amino acid (e.g., arginine, serine, histidine) can be at least one lysine residue selected from the group consisting of

lysine 20, 42, 59, 136, 161,177, 182, 189, 209, 227, 234, 249, 271, 281, 288, 299, 319, 325, 328, 337, 341, 355, 357, 369, 381, 390, 396, 403, 414 and 422 of SEQ ID NO: 816.

The flagellin can be an *E. coli* fliC that includes SEQ ID NO: 814. The flagellin can be a *S. muenchen* that include the includes SEQ ID NO: 813. The flagellin can be a *P. aeruginosa* flagellin that includes SEQ ID NO: 815. The flagellin can be a *Listeria monocytogenes* flagellin that includes SEQ ID NO: 820.

The compositions of the invention can include a flagellin that has lysines substituted in a region adjacent to the motif C of the flagellin, the motif N of the flagellin, both the motif C and the motif N of the flagellin, domain 1 of the flagellin, domain 2 of the flagellin or any combination thereof. Motif C and motif N of flagellin and domain 1 and domain 2 of the flagellin can be involved in activation of TLR5 by the flagellin (Murthy, *et al.*, *J. Biol. Chem.*, 279:5667-5675 (2004)).

Chemical conjugation of a protein, peptide, or polypeptide to another molecule can be by derivatization of a secondary group, such as a lysine. Certain lysine residues in flagellin are near or in domain 1, the motif C or motif N, motifs can be important in binding of the flagellin to TLR5. For example, lysine residues at amino acids 58, 135, 160 and 410 of SEQ ID NO: 812 may be substituted with at least one member selected from the group consisting of an arginine residue, a serine residue and a histidine residue. Derivatization of such lysine residues to, for example, chemically conjugated antigens to flagellins, may decrease the ability or the binding affinity of the flagellin to TLR5 and, thus, diminish an innate immune response mediated by TLR5. Substitution of at least one lysine residue in a flagellin that may be near to regions of the flagellin that are important in mediating interactions with TLR5 (e.g., motif C, motif N, domain 1) with another amino acid (e.g., arginine, serine, histidine) may preserve or enhance flagellin binding to TLR5. In a particular embodiment, the amino acid substitution is a conservative amino acid substitution with at least one member selected from the group consisting of arginine, serine and histidine. Exemplary commercially available reagents for chemical conjugation are described herein.

Certain lysine residues in flagellin are in the domain (domain 1) and can be important for activation of TLR5. For example, lysine residues at positions 58, 135,

160 and 410 of SEQ ID NO: 812 are in domain 1. Derivatization of such lysine residues to, for example, chemically conjugated antigens, may decrease TLR5 bioactivity and, thus, diminish an innate immune response mediated by TLR5.

Lysine residues that can be substituted can include lysine residues implicated in TLR5 activation. Lysine residues in motif N (amino acids 95-108 of SEQ ID NO: 812) and/or motif C (amino acids 441-449 of SEQ ID NO: 812) can be suitable for substitution. Substitution of certain lysine residue in the flagellin (e.g., lysine at amino acid position 19, 41) with, for example, an arginine, serine or histidine, can maintain binding of the flagellin to TLR5 and leave other lysines available for chemical conjugate to another molecule, such as an antigen (e.g., protein) or another molecule, such as another protein, peptide or polypeptide.

The X-ray crystal structure of the F41 fragment of flagellin from *Salmonella* typhimurium shows the domain structure of flagellin (Samatey, F.A., *et al., Nature 410*:321 (2001)). The full length flagellin protein contains 4 domains, designated as D0, D1, D2 and D3. Three of these domains are shown in the crystal structure because the structure was made with a proteolytic fragment of full length flagellin. The amino acid sequences of *Salmonella typhimurium* flagellin for these regions, numbered relative to SEQ ID NO: 812 are as follows:

D0 contains the regions A1 through A55 and S451 through R494

D1 contains the regions N56 through Q176 and T402 through R450

D2 contains the regions K177 through G189 and A284 through A401

D3 contains the region Y190 though V283

Exemplary lysine residues of SEQ ID NO: 812 suitable for substitution with, for example, arginine, histidine, or serine, can include:

D0 contains 2 lysine residues; K19, K41

D1 contains 4 lysine residues; K58, K135, K160 and K410

D2 contains 14 lysine residues at positions 177, 179, 292, 308, 317, 326, 338, 348, 357, 362, 369, 378, 384, 391

D3 contains 8 lysine residues at positions 203, 215, 221,228, 232, 241, 251, 279

Exemplary lysine residues suitable for substitution include lysines at positions 58, 135, 160 and 410 of SEQ ID NO: 812 (Jacchieri, S. G., *et. al., J.*

Bacteriol. 185:4243 (2003); Donnelly, M.A., *et al., J. Biol. Chem. 277*:40456 (2002)). The sequences were obtained from the Swiss-Prot Protein Knowledgebase located online at http://us.expasy.org/sprot/. Lysine residues that can be modified are indicated with a *.

Exemplary lysine residues of SEQ ID NO: 816 suitable for substitution can include:

D0- with two lysines at positions 20, 42;

D1- with five lysines at positions 59, 136, 161, 414, 422;

D2- with sixteen lysines at positions 177, 182, 189, 299, 319, 325, 328, 337, 341, 355, 357, 369, 381, 390, 396, 403 and

D3- with seven lysines at positions 209, 227, 234, 249, 271, 281, 288.

In a further embodiment, the invention is a composition comprising a flagellin component that is at least a portion of a flagellin, wherein at least one lysine of the flagellin component has been substituted with at least one histidine residue, whereby the flagellin component activates a Toll-like Receptor 5.

In yet another embodiment, the invention is a method of stimulating an immune response in a subject, comprising the step of administering to the subject a composition that includes a flagellin component that is at least a portion of a flagellin, wherein the flagellin component includes at least one cysteine residue and whereby the flagellin component activates a Toll-like Receptor 5.

In another embodiment, the invention is a method of stimulating protective immunity in a subject, comprising the step of administering to the subject a composition that includes a flagellin component that is at least a portion of a flagellin, wherein the flagellin component includes at least one cysteine residue and whereby the flagellin component activates a Toll-like Receptor 5.

In yet another embodiment, the invention is a method of stimulating protective immunity in a subject, comprising the step of administering to the subject a composition that includes a flagellin component that is at least a portion of a flagellin, wherein at least one lysine of the flagellin component has been substituted with at least one arginine, whereby the flagellin component activates a Toll-like Receptor 5.

In still another embodiment, the invention is a method of stimulating an immune response in a subject, comprising the step of administering to the subject a composition that includes a flagellin component that is at least a portion of a flagellin, wherein at least one lysine of the flagellin component has been substituted with at least one serine residue, whereby the flagellin component activates a Toll-like Receptor 5.

In another embodiment, the invention is a method of stimulating a protective immune response in a subject, comprising the step of administering to the subject a composition that includes a flagellin component that is at least a portion of a flagellin, wherein at least one lysine of the flagellin component has been substituted with at least one serine residue, whereby the flagellin component activates a Toll-like Receptor 5.

In an additional embodiment, the invention is a method of stimulating an immune response in a subject, comprising the step of administering to the subject a composition that includes a flagellin component that is at least a portion of a flagellin, wherein at least one lysine of the flagellin component has been substituted with at least one histidine residue, whereby the flagellin component activates a Toll-like Receptor 5.

In a further embodiment, the invention is a method of stimulating protective immunity in a subject, comprising the step of administering to the subject a composition that includes a flagellin component that is at least a portion of a flagellin, wherein at least one lysine of the flagellin component has been substituted with at least one histidine residue, whereby the flagellin component activates a Toll-like Receptor 5.

Another embodiment of the invention is a method of stimulating an immune response in a subject, comprising the step of administering to the subject a composition that includes a Toll-like Receptor agonist component that is at least a portion of a Toll-like Receptor agonist, wherein the Toll-like Receptor agonist component includes at least one cysteine residue in a position where a cysteine residue does not occur in the native Toll-like Receptor agonist, whereby the Toll-like Receptor agonist component activates a Toll-like Receptor agonist.

Another embodiment of the invention is a method of stimulating a protective immunity in a subject, comprising the step of administering to the subject a composition that includes a Toll-like Receptor agonist component that is at least a portion of a Toll-like Receptor agonist, wherein the Toll-like Receptor agonist component includes at least one cysteine residue in a position where a cysteine residue does not occur in the native Toll-like Receptor agonist, whereby the Toll-like Receptor agonist component activates a Toll-like Receptor agonist.

In yet another embodiment, the invention is a composition comprising an antigen component that includes at least a portion of a hemagglutinin maturational cleavage site and an agonist component that includes a Toll-like Receptor agonist (e.g., at least one member selected from the group consisting of a Toll-like Receptor 1 agonist, a Toll-like Receptor 3 agonist, a Toll-like Receptor 5 agonist, a Toll-like Receptor 6 agonist, a Toll-like Receptor 7 agonist and a Toll-like Receptor 9 agonist) wherein the Toll-like Receptor agonist is not a Toll-like Receptor 2 agonist (e.g., the outer membrane protein complex (OMPC), such as OMPC of *Neiserria meningitides*). The hemagglutinin maturational cleavage site can include at least one member selected from the group consisting of an influenza A hemagglutinin maturational cleavage site, an influenza B hemagglutinin maturational cleavage site and an influenza C hemagglutinin maturational cleavage site.

In one embodiment, the antigen component and the agonist component of the compositions of the invention can be components of a fusion protein. In another embodiment, the antigen component is chemically conjugated to the agonist component. The agonist component includes a composition comprising a flagellin component that is at least a portion of a flagellin, wherein the flagellin component includes at least one cysteine residue and whereby the flagellin component activates a Toll-like Receptor 5. The antigen component further includes at least a portion of a matrix 2 protein, which can further include a second agonist component that includes at least a portion of a second Toll-like Receptor agonist (e.g., at least a portion of at least one member selected from the group consisting of a Toll-like Receptor 2 agonist, a Toll-like Receptor 3 agonist, a Toll-like Receptor 4 agonist, a Toll-like Receptor 5 agonist, a Toll-like Receptor 6 agonist, a Toll-like Receptor 7

agonist and a Toll-like Receptor 9 agonist). The matrix-2 protein can be fused to the second agonist component or chemically conjugated to a second agonist component.

An "antigen component," as used herein, refers to a part or the entirety of an antigen.

An "agonist component" as used herein, refers to a part or the entirety of an agonist, such as a Toll-like Receptor agonist.

In an additional embodiment, the invention is a method of stimulating an immune response in a subject, comprising the step of administering to the subject a composition comprising an antigen component that includes at least a portion of a hemagglutinin maturational cleavage site and an agonist component that includes a Toll-like Receptor agonist, wherein the Toll-like Receptor agonist is not a Toll-like Receptor 2 agonist.

In a further embodiment, the invention is a method of stimulating protective immunity in a subject, comprising the step of administering to the subject a composition comprising an antigen component that includes at least a portion of a hemagglutinin maturational cleavage site and an agonist component that includes a Toll-like Receptor agonist, wherein the Toll-like Receptor agonist is not a Toll-like Receptor 2 agonist.

In yet another embodiment, the invention is a composition that includes at least a portion of a HA antigen (e.g., the maturational cleavage site) that is fused (e.g., recombinantly) to or chemically conjugated to at least of portion of a hinge region of a flagellin, a flagellin component or in a flagellin in which the hinge region has been deleted; and at least a portion of an M2e protein that is fused (e.g., recombinantly) to or chemically conjugated to a flagellin or a flagellin component, for example, at the amino- or carboxy-terminus of the flagellin or the flagellin component (See, for example, Figure 77).

In an additional embodiment, the invention includes a protein, peptide polypeptide having at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 98% and at least about 99% sequence identity to the proteins, polypeptides and peptides of the invention.

The percent identity of two amino acid sequences (or two nucleic acid sequences) can be determined by aligning the sequences for optimal comparison purposes (e.g., gaps can be introduced in the sequence of a first sequence). The amino acid sequence or nucleic acid sequences at corresponding positions are then compared, and the percent identity between the two sequences is a function of the number of identical positions shared by the sequences (i.e., % identity = # of identical positions/total # of positions x 100). The length of the protein or nucleic acid encoding can be aligned for comparison purposes is at least 30%, preferably, at least 40%, more preferably, at least 60%, and even more preferably, at least 70%, 75%, 80%, 85%, 90%, 95%, 98%, 99% or 100%, of the length of the reference sequence, for example, the nucleic acid sequence of a protein portion of HA (e.g., HA1-1, HA1-2), fusion protein, antigen, or polypeptide.

The actual comparison of the two sequences can be accomplished by well-known methods, for example, using a mathematical algorithm. A preferred, non-limiting example of such a mathematical algorithm is described in Karlin *et al.* (*Proc. Natl. Acad. Sci. USA, 90*:5873-5877 (1993), the teachings of which are hereby incorporated by reference in its entirety). Such an algorithm is incorporated into the BLASTN and BLASTX programs (version 2.2) as described in Schaffer *et al.* (*Nucleic Acids Res., 29*:2994-3005 (2001), the teachings of which are hereby incorporated by reference in its entirety). When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g., BLASTN; available at the Internet site for the National Center for Biotechnology Information) can be used. In one embodiment, the database searched is a non-redundant (NR) database, and parameters for sequence comparison can be set at: no filters; Expect value of 10; Word Size of 3; the Matrix is BLOSUM62; and Gap Costs have an Existence of 11 and an Extension of 1.

Another mathematical algorithm employed for the comparison of sequences is the algorithm of Myers and Miller, CABIOS (1989), the teachings of which are hereby incorporated by reference in its entirety. Such an algorithm is incorporated into the ALIGN program (version 2.0), which is part of the GCG (Accelrys, San Diego, California) sequence alignment software package. When utilizing the ALIGN program for comparing amino acid sequences, a PAM120 weight residue

table, a gap length penalty of 12, and a gap penalty of 4 is used. Additional algorithms for sequence analysis are known in the art and include ADVANCE and ADAM as described in Torellis and Robotti (Comput. Appl. Biosci., 10: 3-5 (1994), the teachings of which are hereby incorporated by reference in its entirety); and FASTA described in Pearson and Lipman (*Proc. Natl. Acad. Sci USA, 85*: 2444-2448 (1988), the teachings of which are hereby incorporated by reference in its entirety).

The percent identity between two amino acid sequences can also be accomplished using the GAP program in the GCG software package (Accelrys, San Diego, California) using either a Blossom 63 matrix or a PAM250 matrix, and a gap weight of 12, 10, 8, 6, or 4 and a length weight of 2, 3, or 4. In yet another embodiment, the percent identity between two nucleic acid sequences can be accomplished using the GAP program in the GCG software package (Accelrys, San Diego, California), using a gap weight of 50 and a length weight of 3.

The nucleic acid sequence encoding a protein portion of HA, polypeptide or fusion proteins of the invention and polypeptides of the invention can include nucleic acid sequences that hybridize to nucleic acid sequences or complements of nucleic acid sequences of the invention, for example, SEQ ID NOS: 53-58, 64, 68, 71, 72, 73, 76, 78, 80, 84, 87, 104, 107, 110, 111, 112, 115, 116, 117, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 137, 139, 142, 143, 146, 147, 150, 167, 184, 185, 186, 187, 188, 189, 190, 191, and 192); and nucleic acid sequences that encode amino acid sequences and fusion proteins of the invention (e.g., SEQ ID NOS: 89-92, 95, 151-160, 177, 209, 210 and 211) under selective hybridization conditions (e.g., highly stringent hybridization conditions). As used herein, the terms "hybridizes under low stringency," "hybridizes under medium stringency," "hybridizes under high stringency," or "hybridizes under very high stringency conditions," describe conditions for hybridization and washing of the nucleic acid sequences. Guidance for performing hybridization reactions, which can include aqueous and nonaqueous methods, can be found in Aubusel, F.M., *et al.*, Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (2001), the teachings of which are hereby incorporated herein in its entirety.

For applications that require high selectivity, relatively high stringency conditions to form hybrids can be employed. In solutions used for some membrane based hybridizations, addition of an organic solvent, such as formamide, allows the reaction to occur at a lower temperature. High stringency conditions are, for example, relatively low salt and/or high temperature conditions. High stringency are provided by about 0.02 M to about 0.10 M NaCl at temperatures of about 50°C to about 70°C. High stringency conditions allow for limited numbers of mismatches between the two sequences. In order to achieve less stringent conditions, the salt concentration may be increased and/or the temperature may be decreased. Medium stringency conditions are achieved at a salt concentration of about 0.1 to 0.25 M NaCl and a temperature of about 37°C to about 55°C, while low stringency conditions are achieved at a salt concentration of about 0.15 M to about 0.9 M NaCl, and a temperature ranging from about 20°C to about 55°C. Selection of components and conditions for hybridization are well known to those skilled in the art and are reviewed in Ausubel *et al.* (1997, Short Protocols in Molecular Biology, John Wiley & Sons, New York N.Y., Units 2.8-2.11, 3.18-3.19 and 4-64.9).

A "subject," as used herein, can be a mammal, such as a primate or rodent (e.g., rat, mouse). In a particular embodiment, the subject is a human.

An "effective amount," when referring to the amount of a composition and fusion protein of the invention, refers to that amount or dose of the composition and fusion protein, that, when administered to the subject is an amount sufficient for therapeutic efficacy (e.g., an amount sufficient to stimulate an immune response in the subject). The compositions and fusion proteins of the invention can be administered in a single dose or in multiple doses.

The methods of the present invention can be accomplished by the administration of the compositions and fusion proteins of the invention by enteral or parenteral means. Specifically, the route of administration is by oral ingestion (e.g., drink, tablet, capsule form) or intramuscular injection of the composition and fusion protein. Other routes of administration as also encompassed by the present invention including intravenous, intradermal, intraarterial, intraperitoneal, or subcutaneous routes, and nasal administration. Suppositories or transdermal patches can also be employed.

The compositions and fusion proteins of the invention can be administered *ex vivo* to a subject's autologous dendritic cells. Following exposure of the dendritic cells to the composition and fusion protein of the invention, the dendritic cells can be administered to the subject.

The compositions and fusion proteins of the invention can be administered alone or can be coadministered to the patient. Coadminstration is meant to include simultaneous or sequential administration of the composition, fusion protein or polypeptide of the invention individually or in combination. Where the composition and fusion protein are administered individually, the mode of administration can be conducted sufficiently close in time to each other (for example, administration of the composition close in time to administration of the fusion protein) so that the effects on stimulating an immune response in a subject are maximal. It is also envisioned that multiple routes of administration (e.g., intramuscular, oral, transdermal) can be used to administer the compositions and fusion proteins of the invention.

The compositions and fusion proteins of the invention can be administered alone or as admixtures with conventional excipients, for example, pharmaceutically, or physiologically, acceptable organic, or inorganic carrier substances suitable for enteral or parenteral application which do not deleteriously react with the extract. Suitable pharmaceutically acceptable carriers include water, salt solutions (such as Ringer's solution), alcohols, oils, gelatins and carbohydrates such as lactose, amylose or starch, fatty acid esters, hydroxymethycellulose, and polyvinyl pyrolidine. Such preparations can be sterilized and, if desired, mixed with auxillary agents such as lubricants, preservatives, stabilizers, wetting agents, emulsifiers, salts for influencing osmotic pressure, buffers, coloring, and/or aromatic substances and the like which do not deleteriously react with the compositions, fusion proteins or polypeptides of the invention. The preparations can also be combined, when desired, with other active substances to reduce metabolic degradation. The compositions and fusion proteins of the invention can be administered by is oral administration, such as a drink, intramuscular or intraperitoneal injection or intranasal delivery. The compositions and fusion proteins alone, or when combined with an admixture, can be administered in a single or in more than one dose over a

period of time to confer the desired effect (e.g., alleviate prevent viral infection, to alleviate symptoms of virus infection, such as influenza or flaviviral infection).

When parenteral application is needed or desired, particularly suitable admixtures for the compositions and fusion proteins are injectable, sterile solutions, preferably oily or aqueous solutions, as well as suspensions, emulsions, or implants, including suppositories. In particular, carriers for parenteral administration include aqueous solutions of dextrose, saline, pure water, ethanol, glycerol, propylene glycol, peanut oil, sesame oil, polyoxyethylene-block polymers, and the like. Ampules are convenient unit dosages. The compositions, fusion proteins or polypeptides can also be incorporated into liposomes or administered via transdermal pumps or patches. Pharmaceutical admixtures suitable for use in the present invention are well-known to those of skill in the art and are described, for example, in Pharmaceutical Sciences (17th Ed., Mack Pub. Co., Easton, PA) and WO 96/05309 the teachings of which are hereby incorporated by reference.

The compositions and fusion proteins of the invention can be administered to a subject on a support that presents the compositions, fusion proteins and polypeptides of the invention to the immune system of the subject to generate an immune response in the subject. The presentation of the compositions, fusion proteins and polypeptides of the invention would preferably include exposure of antigenic portions of the viral protein to generate antibodies. The components (e.g., PAMP and a viral protein) of the compositions, fusion proteins and polypeptides of the invention are in close physical proximity to one another on the support. The compositions and fusion proteins of the invention can be attached to the support by covalent or noncovalent attachment. Preferably, the support is biocompatible. "Biocompatible," as used herein, means that the support does not generate an immune response in the subject (e.g., the production of antibodies). The support can be a biodegradable substrate carrier, such as a polymer bead or a liposome. The support can further include alum or other suitable adjuvants. The support can be a virus (e.g., adenovirus, poxvirus, alphavirus), bacteria (e.g., *Salmonella*) or a nucleic acid (e.g., plasmid DNA).

The dosage and frequency (single or multiple doses) administered to a subject can vary depending upon a variety of factors, including prior exposure to an

antigen, a viral protein, the duration of viral infection, prior treatment of the viral infection, the route of administration of the composition, fusion protein or polypeptide; size, age, sex, health, body weight, body mass index, and diet of the subject; nature and extent of symptoms of viral exposure, viral infection and the particular viral responsible for the infection (e.g., a flavivirus, influenza virus), or treatment or infection of an other antigen, such as an influenza antigen, kind of concurrent treatment, complications from the viral exposure, viral infection or exposure or other health-related problems. Other therapeutic regimens or agents can be used in conjunction with the methods and compositions, fusion proteins or polypeptides of the present invention. For example, the administration of the compositions and fusion proteins can be accompanied by other viral therapeutics or use of agents to treat the symptoms of a condition associated with or consequent to exposure to the antigen, such as flavivirus infection (e.g., high fever, numbness, DHF, meningoencephalitis) or influenza infection, for example. Adjustment and manipulation of established dosages (e.g., frequency and duration) are well within the ability of those skilled in the art.

Influenza viruses are single-stranded RNA viruses that belong to the viral family *Orthomyxoviridae*. Influenza viruses are divided into three types (A, B, C) determined by the antigenic differences in ribonucleoprotein (RNP) and matrix (M) antigens of the viruses. Influenza A virus naturally infects humans and several other mammalian species, including swine and horses, and a wide variety of avian species, and causes epidemics and pandemics in the human population. Influenza B virus appears to naturally infect only humans and can cause epidemics in humans. Influenza C virus has been isolated from humans and swine, but generally does not occur in epidemics and usually results in mild disease in humans.

Mature influenza virions are enveloped with a pleomorphic structure ranging in diameter from 80 to 120 nm. The single-stranded RNA genome is closely associated with a helical nucleoprotein and is present in seven (influenza C) or eight (influenza A and B) separate segments of ribonucleoprotein (RNP), each of which has to be present for successful replication of the virus. The segmented genome is enclosed within an outer lipoprotein envelope. Matrix protein 1 (MP1 or also

referred to herein as "M1") lines the inside of the outer lipoprotein envelope and is bound to the RNP.

Hemagglutinin (HA) is a surface glycoprotein on a virus (e.g., an influenza virus) that is responsible for binding to N-AcetylNeuraminic Acid (NeuNAc; also referred to herein as "sialic acid") on host cells and subsequent fusion of viral and host membranes. HA acquired its name by virtue of its ability to cause red blood cells to clump, or agglutinate. Influenza HA is a trimer consisting of the three monomeric (HA0) subunits. HA performs two critical functions during the infection process: binding to a cell surface sialyloligosaccharide receptor and fusion of virus and host cell membrane. Following binding of the HA trimer to the plasma membrane of a host cell, the host cell membrane engulfs the virus in an endosome and attempts to digest the contents of the endosome by acidifying its interior and transferring it to a lysosome in the host cell. However, the acidic environment of the lysosome destabilizes HA, resulting in partial unfolding of HA0 which exposes a protease-sensitive site (the maturaional cleaveage site) that is cleaved by a host protease to form HA1 and HA2 subunits which are connected by a single disulfide bond (Wiley, D. C., *et al.*, *Annu. Rev. Biochem. 56*:365-394 (1987)). Cleavage occurs at a specific amino acid residue and generates a hydrophobic amino terminus for the HA2 subunit. This hydrophobic terminus of HA2 mediates fusion between the viral envelope and the endosomal membrane of the host cell and releases the contents of the virion into the cytoplasm of the cell, a process known as uncoating. Thus, cleavage of the HA polypeptide is a requirement for infectivity.

The crystal structure of several viral hemagglutinins has been determined (see, for example, Wilson, I.A., *et al.*, *Nature 289*:366-373 (1981); Chen, J., *et al.*, *Cell 95*:409-417 (1998); Ha, Y., *et al.*, *The EMBO Journal 21*: 865-875 (2002); Russell, R.J., *et al.*, *Virology 325*:287-296 (2004); and Cox, N.J., *et al.*, In: Toply and Wilson's Microbiology and Microbial Infections, eds. B.W.J. Mathy, *et al.*, Vol. 1 (9th ed.) New York, NY, Oxford Univ. Press, Ch. 32, p. 634 (1998)). X-ray crystallographic structures show that HA is folded into two structural components or domains – a globular head and a fibrous stalk (see, for example, Figure 1). The globular head includes HA1, including that part of HA1 that binds to sialic acid (also referred to as the "receptor binding site or domain" or "sialic acid binding site or

domain"), and antiparallel β-sheets. The fibrous stalk is more proximal to the viral membrane and consists of all of HA2 and part of HA1, including the cleavage site between HA1 and HA2.

There are fifteen known subtypes of Influenza A HA (H1-H15) that share between about 40 to about 60% sequence identity (World Health Organization BULL. World Health Organ., 58:585-591 (1980)). Influenza viruses containing all 15 HA subtypes have been isolated from avian species (H5, H7, and H9), equine (H3 and H7), seals (H3, H4 and H7), whales (H1 and H13) and swine (H1, H3, and H9). Subtypes of influenza A virus are generally named according to the particular antigenic determinants of HA (H, 15 major types) and neuraminidase (N, about 9 major types). For example, subtypes include influenza A (H2N1), A(H3N2), A(H5N1), A(H7N2), A(H9N2), A(H1/H0), A(H3/H0) and A(H5/H0). In the last century, three subtypes of influenza A resulted in pandemics: H1 in 1918 and 1977; H2 in 1957 and H3 in 1968. In 1997, an H5 avian virus and in 1999, an H9 virus resulted in outbreaks of respiratory disease in Hong Kong. HA from influenza type B viruses have been isolated from humans and seals and are not divided into subtypes.

A host infected with influenza can mount an antibody response to the globular head of HA that protects that host from subsequent infection with the same strain of virus by blocking the interaction between HA and the host cell, i.e., neutralizing the infectivity of the virus. Due to the low fidelity and high rate of influenza RNA replication, the virus is constantly experiencing minor mutations in the HA gene that preserve the globular head structure and host cell interaction, but may allow progeny virus to escape immune surveillance. These point mutations are referred to as "antigenic drift." In addition, if a single host is simultaneously infected with two different strains of influenza A, a new subtype of virus may emerge as a result of reassortment, or the exchange of the RNA segments, or genes, between different strains of influenza A viruses. The viruses emerging from reassortment present the human immune system with a new antigenic experience that usually results in high morbidity and mortality. This type of drastic antigenic change is known as "antigenic shift." Since type B influenza viruses circulate

almost exclusively in humans, these viruses cannot undergo reassortment with animal strains and, thus, are changed only by antigenic drift.

Immunity to HA can reduce the likelihood of infection and severity of disease if infection does occur. HA is an important antigenic target and the efficacy of vaccines depends on the antigenic match between the vaccine strain and the circulating strain. Since the hemagglutinin protein readily undergoes antigenic shift and drift in order to evade the host's immune defense, traditional vaccines must be based on currently circulating influenza strains and annually updated. Annual updates of influenza vaccines are not only costly they also require significant amounts of production time and manufacturing infrastructure. A vaccine composition based on invariant regions of the virus may provide broadly cross-reactive protection.

In contrast to the globular head of HA, changes in amino acid residues surrounding the maturational cleavage site of HA are limited due to functional constraints. Amino acid residues surrounding the HA maturational cleavage site influence recognition and therefore cleavability of the site by the host protease. Since the virus does not code for the protease, changes in the amino acid residues surrounding the maturational cleavage site are restricted. As a consequence a peptide of about 20 amino acids spanning the maturational cleavage site remains genetically stable across influenza viruses of the same HA subtype (WO 2004/080403; Bianchi, *et al. J Virol 79:*7380-7388 (2005)) or as branched peptides (Horvath, *et al Immunol Letters* 60:127-136(1998), Nagy, *et al Scand J Immunol 40*:281-291 (1994)).

A second highly conserved antigen of influenza A is the ectodomain of the matrix 2 protein (M2e). M2 is a 97-amino acid protein expressed at low levels in mature virions and much higher levels on infected cells. The M2 protein forms a homotetramer that functions as an ion channel which is critical to the replication of the virus, thus, mutations in M2e are not as well tolerated as mutations in HA. The 24-amino acid ectodomain (M2e) is highly conserved across multiple influenza A strains. In mammals, M2e is poorly immunogenic in its native form. Antibodies to M2e can confer passive protection in animal models of influenza A infection (Treanor, J.J., *et al., J Virol 64*:1375 (1990); Liu, W.P., *et al., Immunol Lett 93:131*

(2004)), not by neutralizing the virus and preventing infectivity, but by killing infected cells and disrupting the viral life cycle (Zebedee, S.L., *et al., J. Virol. 62*:2762 (1998); Jegerlehner, A.N., *et. al., J Immunol 172*:5598 (2004)), which may be by antibody-dependent NK cell activity (Jegerlehner, A.N., *et al., J Immunol 172*:5598 (2004)). Composition that include M2e proteins may limit the severity of influenza A disease while allowing the host immune response to develop adaptive immunity to the dominant neutralizing influenza antigen, HA.

Strategies to manage infection and illness consequent to influenza viral infection have not changed significantly in the past four decades. Due to the seasonal nature of influenza disease, the distinct types of influenza virus (A and B) that threaten the human population, and the genetic instability of each type, it is necessary to reformulate a multivalent vaccine each year, based on epidemiological prediction of strains likely to be circulating in the human population in the upcoming flu season. The vaccine is produced from stocks of selected prototype viral strains grown in embryonated chicken eggs. Thus, the current strategy has several limitations that include: (a) dependence on uncertain prediction of circulating strains; (b) dependence on the ability to grow the appropriate strains in chicken eggs; (c) the egg-based production system carries risks of product contamination; (d) the product produced in eggs cannot be used in individuals with egg allergies; and (e) a significant risk that the typical multivalent vaccine will not confer protection against a pandemic strain of virus to which the human population has no pre-exisiting immunity.

The dominant protective component of the currently-available influenza vaccine is the viral hemagglutinin (HA). A more effective vaccine may include not only strain-specific HA, but cross-protective antigens, such as M2e and maturational cleavage site. A vaccine production process that is more reliable, economical, and scaleable than the current egg-based method is also preferable. The compositions, fusion proteins and polypeptides of the invention provide compositions that include HA, M2e, and maturational cleavage site of influenza proteins, which can stimulate an immune response, specifically, a protective immune response to several influenza antigens in subject.

The teachings of all patents, published applications and references cited herein are incorporated by reference in their entirety.

EXEMPLIFICATION

EXAMPLE 1: DESIGN OF PORTIONS OF A NATURALLY OCCURRING HEMAGGLUTININ OF INFLUENZA A

Materials and Methods

Design of the HA1-1, 1-2 and 1-3 globular head constructs for A/Puerto Rico/8/34 (H1N1). Influenza strain A/Puerto Rico/8/34 (PR8) is a well-characterized mouse-adapted strain of influenza A virus. The published crystal structure of the mature PR8 HA (SEQ ID NO: 1) (Gamblin et al. 2005. *Science* 303:1838-1842; PDB accession number 1RU7) was used in combination with the Molecular Modeling Database to determine the boundaries for three PR8 globular head constructs. The complete list of solved crystal structures for influenza A hemagglutinin molecules can be found in the Protein Data Bank (PDB) or National Center for Biotechnology Information (NCBI) Structure website (http://www.ncbi.nlm.nih.gov/Structure).

The three dimensional crystal structure for the PR8 HA was viewed using the Cn3D program within the NCBI website. The trimeric HA molecule has the shape of a mushroom. The globular head refers to the portion of the hemagglutinin molecule at the top that resembles the mushroom head, and the coiled-coil stalk refers to the bottom part of the molecule that resembles the stalk of the mushroom (Figure 1). The globular head contains the substrate binding site that binds to sialic acid on the cell surface, and therefore is important for viral entry. In addition, the majority of neutralizing antibodies epitopes are located near and around this receptor binding site, making the globular head a good target for protective vaccines. The globular head contains most of the HA1 peptide including residues numbered from E51 to K327 in PR8 for example.

The monomeric structure of PR8 chain A, which encompasses the globular head, was used to guide the design of the PR8 HA1-1, 1-2 and 1-3 constructs (Figure 1). The HA constructs were designed with domain boundaries within the globular head so that when the molecule is expressed in a host cell, the encoded

protein it may spontaneously fold or refolded in vitro to mimic the native conformation. A structural domain (also referred to herein as "domains") within a protein is an element of overall structure that is self-stabilizing and often folds independently of the rest of the protein chain. Most domains can be classified into folds. Many domains are not unique to the proteins produced by one gene or one gene family but instead appear in a variety of proteins. Domains often are named and singled out because they play an important role in the biological function of the protein to which they belong; for example, the substrate binding domain of hemagglutinin can participate in binding to the substrate. Because domains can be self-stabilizing, domains can be swapped between one protein and another or associated to other carrier proteins by genetic engineering to make chimera proteins. A domain may be composed of none, one, or many structural motifs. In the cases of influenza hemagglutinin construct design, many structural motifs are preserved. The boundary selection is guided by known crystal structures. In the case of PR8, residues from E51 to K327 (SEQ ID NO: 1) fold into a compact structure that distinguishes from the rest of the HA molecules, thus is the region of focus to design HA constructs.

Design of the HA1-1, 1-2 and 1-3 globular head constructs for A/Viet Nam/1203/2004 (H5N1). The HA structure used as a reference in the design of the Viet Nam/1203/2004 (H5N1) globular head constructs is described in Stevens et al, 2006, *Science* 312:404-410 (MMDB number 38730; PDB accession number 2FK0). As described for the PR8 globular head constructs above, the published crystal structure of the mature A/Viet Nam/1203/2004 HA (SEQ ID NO: 2) was used in combination with the Molecular Modeling Database to determine the domain boundaries for three Viet Nam globular head constructs. The same structural criteria for preservation of secondary and tertiary structure for the structural domain were applied to the design of the Viet Nam constructs. While different hemagglutinin molecules can differ in the nature or number of residues that comprise the hemagglutinin molecule; the overall structure is remarkably similar. Thus, design of the domain boundaries for the Viet Nam globular head constructs required placement of the boundaries at structurally equivalent but numerically different positions with in the HA molecule.

Design of the HA1-1, 1-2 and 1-3 globular head constructs for A/Indonesia/5/2005 (H5N1). The crystal structure for the Indonesia HA has not been solved. In general, when the crystal structure for a given HA molecule has not been solved, the available structure with the highest sequence identity can be used to guide the construct design. In the case of the design of the Indonesia globular head construct, the closest structure available is the A/Viet Nam/1203/2004 (Stevens et al. 2006. *Science* 312:404-410; MMDB number 38730; PDB Accession number 2FK0), which is of the same subtype as A/Indonesia/5/2005.

To design globular head constructs for this HA, the primary sequence of the Indonesia HA (SEQ ID NO: 3) was first aligned with A/Viet Nam/1203/2004 HA (H5VN; SEQ ID NO: 2) molecule (primary sequence identity: 96.13%). The primary sequence alignment for A/Viet Nam/1203/2004 HA (H5VN; SEQ ID NO: 2) and A/Indonesian/5/2005 (H5IN; SEQ ID NO: 3) was conducted using CLUSTALW and is shown below, where (*) asterisk = identity, (:) colon = conservative substitution; (.) period = weakly conservative substitution; and (space) = divergent substitution.

```
                                                    50

H5IN   MEKIVLLLAIVSLVKSDQICIGYHANNSTEQVDTIMEKNVTVTHAQDILE
H5VN   MEKIVLLFAIVSLVKSDQICIGYHANNSTEQVDTIMEKNVTVTHAQDILE
       ********:**************************************************

                                                   100

H5IN   KTHNGKLCDLDGVKPLILRDCSVAGWLLGNPMCDEFINVPEWSYIVEKAN
H5VN   KKHNGKLCDLDGVKPLILRDCSVAGWLLGNPMCDEFINVPEWSYIVEKAN
       *.**************************************************

                                                   150

H5IN   PTNDLCYPGSFNDYEELKHLLSRINHFEKIQIIPKSSWSDHEASSGVSSA
H5VN   PVNDLCYPGDFNDYEELKHLLSRINHFEKIQIIPKSSWSSHEASLGVSSA
       *.*******.*********************************.**** *****

                                                   200

H5IN   CPYLGSPSFFRNVVWLIKKNSTYPTIKKSYNNTNQEDLLVLWGIHHPNDA
H5VN   CPYQGKSSFFRNVVWLIKKNSTYPTIKRSYNNTNQEDLLVLWGIHHPNDA
       *** *..**********************:**********************

                                                   250

H5IN   AEQTRLYQNPTTYISIGTSTLNQRLVPKIATRSKVNGQSGRMEFFWTILK
H5VN   AEQTKLYQNPTTYISVGTSTLNQRLVPRIATRSKVNGQSGRMEFFWTILK
       ****:**********:***********:**********************
```

```
                                                                300
H5IN   PNDAINFESNGNFIAPEYAYKIVKKGDSAIMKSELEYGNCNTKCQTPMGA
H5VN   PNDAINFESNGNFIAPEYAYKIVKKGDSTIMKSELEYGNCNTKCQTPMGA
       ****************************:*********************

                                                                350
H5IN   INSSMPFHNIHPLTIGECPKYVKSNRLVLATGLRNSPQRESRRKKRGLFG
H5VN   INSSMPFHNIHPLTIGECPKYVKSNRLVLATGLRNSPQRERRRKKRGLFG
       ***************************************** *********

                                                                400
H5IN   AIAGFIEGGWQGMVDGWYGYHHSNEQGSGYAADKESTQKAIDGVTNKVNS
H5VN   AIAGFIEGGWQGMVDGWYGYHHSNEQGSGYAADKESTQKAIDGVTNKVNS
       ************************************************

                                                                450
H5IN   IIDKMNTQFEAVGREFNNLERRIENLNKKMEDGFLDVWTYNAELLVLMEN
H5VN   IIDKMNTQFEAVGREFNNLERRIENLNKKMEDGFLDVWTYNAELLVLMEN
       ************************************************

                                                                500
H5IN   ERTLDFHDSNVKNLYDKVRLQLRDNAKELGNGCFEFYHKCDNECMESIRN
H5VN   ERTLDFHDSNVKNLYDKVRLQLRDNAKELGNGCFEFYHKCDNECMESVRN
       ********************************************:***

                                                                550
H5IN   GTYNYPQYSEEARLKREEISGVKLESIGTYQILSIYSTVASSLALAIMMA
H5VN   GTYDYPQYSEEARLKREEISGVKLESIGIYQILSIYSTVASSLALAIMVA
       ***:************************ ********************:*

                                                                568
H5IN   GLSLWMCSNGSLQCRICI (SEQ ID NO:3)
H5VN   GLSLWMCSNGSLQCR---(SEQ ID NO:2)
       ***************
```

Design of the HA1-1, 1-2 and 1-3 globular head constructs for A/New Caledonia/20/1999 (H1N1). The crystal structure for the New Caledonia HA has not been solved. To design the globular head constructs for this HA, the primary sequence of the New Caledonia HA (H1NC; SEQ ID NO: 4) was first aligned with two closely related HA molecules of the same subtype (primary sequence homology >85%) for which the structures have been resolved, specifically the H1N1 1918 virus (same as A/South Carolina/1/18) (SEQ ID NO: 5) and the A/Puerto Rico/8/34 virus (H1PR8; SEQ ID NO: 1) (Gamblin, et al., Science 303, 1838-42 (2004). A/South Carolina/1/18: MMDB number 26943, PDB accession number: 1RUZ and PR8, MMDB number: 26941 and PDB accession number: 1RU7). The primary sequence alignment was conducted using CLUSTALW and is shown below, where (*) asterisk = identity, (:) colon = conservative substitution; (.) period = weakly conservative substitution; and (space) = divergent substitution.

```
                                                                    50
H1NC   MKAKLLVLLCTFTATYADTICIGYHANNSTDTVDTVLEKNVTVTHSVNLL
H1PR8  MKANLLVLLSALAAADADTICIGYHANNSTDTVDTVLEKNVTVTHSVNLL
        ***:*****.:::*: ************************************

                                                                   100
H1NC   EDSHNGKLCLLKGIAPLQLGNCSVAGWILGNPECELLISKESWSYIVETP
H1PR8  EDSHNGKLCRLKGIAPLQLGKCNIAGWLLGNPECDPLLPVRSWSYIVETP
        ********* **********:*.:***:******: *:. .*********

                                                                   150
H1NC   NPENGTCYPGYFADYEELREQLSSVSSFERFEIFPKESSWPNHTVTGVSA
H1PR8  NSENGICYPGDFIDYEELREQLSSVSSFERFEIFPKESSWPNHNTNGVTA
        *.*** **** * ********************************...**:*

                                                                   200
H1NC   SCSHNGKSSFYRNLLWLTGKNGLYPNLSKSYVNNKEKEVLVLWGVHHPPN
H1PR8  ACSHEGKSSFYRNLLWLTEKEGSYPKLKNSYVNKKGKEVLVLWGIHHPPN
        :***:************** *:* **:*.:****:* ********:*****

                                                                   250
H1NC   IGNQRALYHTENAYVSVVSSHYSRRFTPEIAKRPKVRDQEGRINYYWTLL
H1PR8  SKEQQNIYQNENAYVSVVTSNYNRRFTPEIAERPKVRDQAGRMNYYWTLL
          :*: :*:.********:*:*.********:******* **:*******

                                                                   300
H1NC   EPGDTIIFEANGNLIAPWYAFALSRGFGSGIITSNAPMDECDAKCQTPQG
H1PR8  KPGDTIIFEANGNLIAPMYAFALSRGFGSGIITSNASMHECNTKCQTPLG
        :**************** ******************.*.**::***** *

                                                                   350
H1NC   AINSSLPFQNVHPVTIGECPKYVRSAKLRMVTGLRNIPSIQSRGLFGAIA
H1PR8  AINSSLPYQNIHPVTIGECPKYVRSAKLRMVTGLRNTPSIQSRGLFGAIA
        *******:**:************************* *************

                                                                   400
H1NC   GFIEGGWTGMVDGWYGYHHQNEQGSGYAADQKSTQNAINGITNKVNSVIE
H1PR8  GFIEGGWTGMIDGWYGYHHQNEQGSGYAADQKSTQNAINGITNKVNTVIE
        **********:*************************************:***

                                                                   450
H1NC   KMNTQFTAVGKEFNKLERRMENLNKKVDDGFLDIWTYNAELLVLLENERT
H1PR8  KMNIQFTAVGKEFNKLEKRMENLNKKVDDGFLDIWTYNAELLVLLENERT
        *** *************:************************************

                                                                   500
H1NC   LDFHDSNVKNLYEKVKSQLKNNAKEIGNGC--------------------
H1PR8  LDFHDSNVKNLYEKVKSQLKNNAKEIGNGCFEFYHKCDNECMESVRNGTY
        *******************************

                                                                   550
H1NC   ----------------------------------------------------
H1PR8  DYPKYSEESKLNREKVDGVKLESMGIYQILAIYSTVASSLVLLVSLGAIS

                                                                   565
H1NC   ---------------  (SEQ ID NO: 4)
H1PR8  FWMCSNGSLQCRICI  (SEQ ID NO: 1)
```

Design of the HA1-1, 1-2 and 1-3 globular head constructs for A/Wisconsin/67/2005 (H3N2). The crystal structure for the Wisconsin HA has not been solved. To design the globular head constructs for this HA, the primary sequence of the Wisconsin HA (H3Wis; SEQ ID NO: 6) was first aligned with a closely related reference HA molecule of the same subtype (primary sequence identity: 81.16%) for which the structure had been resolved, specifically influenza A/X31 subtype H3N2 (H3X31; SEQ ID NO: 7) (PDB accession number: 1VIU). The Wisconsin HA sequence was aligned with the X31 HA using CLUSTAL W and is shown below, where (*) asterisk = identity, (:) colon = conservative substitution; (.) period = weakly conservative substitution; and (space) = divergent substitution. Each amino acid in X31 was found to have a corresponding match in A/Wisconsin/67/2005 sequence. The domain boundaries the X31 structure were then used to identify the domain boundaries in the A/Wisconsin/67/2005.

```
                                                                     50
H3Wis QKLPGNDNSTATLCLGHHAVPNGTIVKTITNDQIEVTNATELVQSSSTGG
H3X31 QDLPGNDNSTATLCLGHHAVPNGTLVKTITDDQIEVTNATELVQSSSTGK
      *.************************:*****:******************

                                                                    100
H3Wis ICDSPHQILDGENCTLIDALLGDPQCDGFQNKKWDLFVERSKAYSNCYPY
H3X31 ICNNPHRILDGIDCTLIDALLGDPHCDVFQNETWDLFVERSKAFSNCYPY
      **:.**:**** :*********:** ***:.**********:******

                                                                    150
H3Wis DVPDYASLRSLVASSGTLEFNDESFNWTGVTQNGTSSACKRRSNNSFFSR
H3X31 DVPDYASLRSLVASSGTLEFITEGFTWTGVIQNGGSNACKRGPGSGFFSR
      ********************  *.*.**** *** *.**** ....****

                                                                    200
H3Wis LNWLTHLKFKYPALNVTMPNNEKFDKLYIWGVHHPGTDNDQIFLHAQASG
H3X31 LNWLTKSGSTYPVLNVTMPNNDNFDKLYIWGIHHPSTNQEQTSLYVQASG
      *****:    .**.********::********:***.*:::*  *:.****

                                                                    250
H3Wis RITVSTKRSQQTVIPNIGSRPRIRNIPSRISIYWTIVKPGDILLINSTGN
H3X31 RVTVSTRRSQQTIIPNIGSRPWVRGLSSRISIYWTIVKPGDVLVINSNGN
      *:****:*****:******** :*.:.***************:*:***.**

                                                                    300
H3Wis LIAPRGYFKIRSGKSSIMRSDAPIGKCNSECITPNGSIPNDKPFQNVNRI
H3X31 LIAPRGYFKMRTGKSSIMRSDAPIDTCISECITPNGSIPNDKPFQNVNKI
      *********:*:************..* ********************:*

                                                                    329
H3Wis TYGACPRYVKQNTLKLATGMRNVPEKQTR (SEQ ID NO: 6)
H3X31 TYGACPKYVKQNTLKLATGMRNVPEKQT- (SEQ ID NO: 7)
      ******:**********************
```

RESULTS

Depiction of domain boundaries for PR8 HA constructs. The selected boundary domains are highlighted in the sequence below (SEQ ID NO: 1) as follows: HA1-1 boundaries are single-underlined (S53-R324 of SEQ ID NO: 1); HA1-2 boundaries are double-underlined (K62-S284 of SEQ ID NO: 1); HA1-3 boundaries are bold-underlined (N101-G276 of SEQ ID NO: 1). Detailed descriptions of each subunit design and boundary domains are given below.

```
                                                                  60
MKANLLVLLSALAAADADTICIGYHANNSTDTVDTVLEKNVTVTHSVNLLEDSHNGKLCR
                                                                 120
LKGIAPLQLGKCNIAGWLLGNPECDPLLPVRSWSYIVETPNSENGICYPGDFIDYEELRE
                                                                 180
QLSSVSSFERFEIFPKESSWPNHNTNGVTAACSHEGKSSFYRNLLWLTEKEGSYPKLKNS
                                                                 240
YVNKKGKEVLVLWGIHHPPNSKEQQNLYQNENAYVSVVTSNYNRRFTPEIAERPKVRDQA
                                                                 300
GRMNYYWTLLKPGDTIIFEANGNLIAPMYAFALSRGFGSGIITSNASMHECNTKCQTPLG
                                                                 360
AINSSLPYQNIHPVTIGECPKYVRSAKLRMVTGLRNIPSIQSRGLFGAIAGFIEGGWTGM
                                                                 420
IDGWYGYHHQNEQGSGYAADQKSTQNAINGITNKVNTVIEKMNIQFTAVGKEFNKLEKRM
                                                                 480
ENLNKKVDDGFLDIWTYNAELLVLLENERTLDFHDSNVKNLYEKVKSQLKNNAKEIGNGC
                                                                 540
FEFYHKCDNECMESVRNGTYDYPKYSEESKLNREKVDGVKLESMGIYQILAIYSTVASSL
                                                                 565
VLLVSLGAISFWMCSNGSLQCRICI (SEQ ID NO: 1)
```

PR/8 HA1-1 construct (SEQ ID NO: 8). For this construct, four (4) conserved disulfide bonds were preserved by making the amino-terminal truncation before the Cysteine at position 59 of SEQ ID NO: 1 and the carboxyl-terminal truncation after the Cysteine at position 319 of SEQ ID NO: 1. The preserved disulfide bonds are listed as the following: C59-C291; C72-C84; C107-C152 and C295-C319 of SEQ ID NO: 1. Using the crystal structure cited above and the Molecular Modeling Database (MMDB) to identify tertiary and secondary structure, the amino terminal truncation was placed at the Serine at position 53 of SEQ ID NO: 1, which is immediately adjacent to the coil turn formed by residues 50-52 of SEQ ID NO: 1 (LED). The truncation completely eliminated the preceding β-strand formed by residues 44-49 of SEQ ID NO: 1 [THSVNL (SEQ ID NO: 214)]. The strand spanning residues 53-58 of SEQ ID NO: 1 [SHNGKL (SEQ ID NO: 215)] and the entire β-strand associated with residues 59-62 of SEQ ID NO: 1 [CRLK

(SEQ ID NO: 216)] were preserved. Although in the structure residues 53-58 of SEQ ID NO: 1 [SHNGKL (SEQ ID NO: 215)] are defined as a random coil, these residues mimic a β-strand and thus serve to further stabilize the β-pleated sheets defined by residues 307-310 of SEQ ID NO: 1 [PYQN, SEQ ID NO: 217] and residues 320-323 of SEQ ID NO: 1 [PKYV, SEQ ID NO: 218]. The carboxyl-terminal truncation was made in the random coil sequence spanning residues 324-327 of SEQ ID NO: 1 [RSAK, SEQ ID NO: 219] at the Arginine at position 324 of SEQ ID NO: 1 such that the carboxyl-terminal tail which interacts with HA2 was completely eliminated.

PR/8 HA1-2 construct (SEQ ID NO: 9). For this construct two (2) conserved disulfide bonds (C72-C84 and C107-C152 of SEQ ID NO: 1) were preserved by making the amino-terminal truncation before the Cysteine at position 72 of SEQ ID NO: 1 and the carboxyl-terminal truncation after the Cysteine at position 152 of SEQ ID NO: 1. Using the crystal structure cited above and the MMD database to identify tertiary and secondary structure, the amino-terminal and the carboxyl-terminal truncations were placed in the loops [KGI, residues 62-64 of SEQ ID NO: 1 and NASMHE (SEQ ID NO: 220), residues 285-290, of SEQ ID NO: 1] that connect two distinct sets of β-pleated sheets. Truncation of the molecule within these loops preserves the surrounding secondary structure which in turn preserves the tertiary structure of the globular head.

The amino-terminal truncation was made at the Lysine at position 62 of SEQ ID NO: 1, such that the β-strand comprising APLQLG (SEQ ID NO: 221) at residues 65-70 of SEQ ID NO: 1 of the membrane-distal set of β-strands remained wholly intact, while the preceding β-strand [CRLK, SEQ ID NO: 216, residues 59-62 of SEQ ID NO: 1] that is more membrane-proximal was largely eliminated. The carboxyl-terminal truncation was made at the Serine at position 284 of SEQ ID NO: 1 such that the β-strand [GIITS, SEQ ID NO: 222, residues 280-284 of SEQ ID NO: 1] of the distal set of β-strands remains wholly or largely intact and the subsequent β-strand [CNTK, SEQ ID NO: 223, residues 291-294 of SEQ ID NO: 1] of the proximal set of β-stands is wholly or largely eliminated. Preservation of the distal set of β-strands was the main determinant in selecting the domain boundaries for the HA1-2 construct. The distal β-pleated sheet, which is comprised of β-strands of

APLQLG (SEQ ID NO: 221) [residues 65-70 of SEQ ID NO: 1], SYIVET (SEQ ID NO: 224) [residues 94-99 of SEQ ID NO: 1], and GIITS (SEQ ID NO: 222) [residues 280-284 of SEQ ID NO: 1], serves as the stabilizing secondary structural element that ensures a compact domain structure.

PR/8 HA1-3 construct (SEQ ID NO: 10). For this construct one (1) conserved disulfide bond (C107-C152 of SEQ ID NO: 1) was preserved by making the amino-terminal truncation before the Cysteine at position 107 of SEQ ID NO: 1 and the carboxyl terminal truncation after the Cysteine at position 152 of SEQ ID NO: 1. Using the crystal structure cited above and the MMDB database to identify tertiary and secondary structure the amino-terminal truncation was made at the Asparagine at position 101 of SEQ ID NO: 1 such that the preceding α-helix [NIAGWLLG, SEQ ID NO: 225, residues 73-80 of SEQ ID NO: 1] and β-strand [SYIVET, SEQ ID NO: 224, residues 94-99 of SEQ ID NO: 1] were completely eliminated while the subsequent β-stand [PGDFI, SEQ ID NO: 226, residues 109-113 of SEQ ID NO: 1] was preserved along with the strand NSENGICY, SEQ ID NO: 227 [residues 101-108 of SEQ ID NO: 1] that includes the conserved Cysteine at position 107 of SEQ ID NO: 1.

The carboxyl-terminal truncation was made near the end of the β-strand [AFALSRGF, SEQ ID NO: 228, residues 270-277 of SEQ ID NO: 1]. The last amino acid of this strand, position 277 of SEQ ID NO: 1, is a Phenylalanine, which was eliminated so as not to expose this hydrophobic residue. The carboxyl-terminal truncation was made at the Glycine at position 276 of SEQ ID NO: 1 which is in between two secondary structures. Truncation at this position keeps the preceding β-strand [AFALSRG, SEQ ID NO: 229, residues 270-276 of SEQ ID NO: 1] largely intact and completely eliminates the subsequent β-strand [GIITS, SEQ ID NO: 222, residues 280-284 of SEQ ID NO: 1].

Depiction of domain boundaries for VN04 HA construct. The selected boundary domains are highlighted in the sequence below (SEQ ID NO: 2) as follows: HA1-1 boundaries are single-underlined (E50-K323 of SEQ ID NO: 2); HA1-2 boundaries are double-underlined (G62-E284 of SEQ ID NO: 2); HA1-3 boundaries are bold-underlined (N103-G276 of SEQ ID NO: 2). Detailed descriptions of each subunit design and boundary domains are given below.

```
                                                                  60
MEKIVLLFAIVSLVKSDQICIGYHANNSTEQVDTIMEKNVTVTHAQDILEKKHNGKLCDL
                                                                 120
DGVKPLILRDCSVAGWLLGNPMCDEFINVPEWSYIVEKANPVNDLCYPGDFNDYEELKHL
                                                                 180
LSRINHFEKIQIIPKSSWSSHEASLGVSSACPYQGKSSFFRNVVWLIKKNSTYPTIKRSY
                                                                 240
NNTNQEDLLVLWGIHHPNDAAEQTKLYQNPTTYISVGTSTLNQRLVPRIATRSKVNGQSG
                                                                 300
RMEFFWTILKPNDAINFESNGNFIAPEYAYKIVKKGDSTIMKSELEYGNCNTKCQTPMGA
                                                                 360
INSSMPFHNIHPLTIGECPKYVKSNRLVLATGLRNSPQRERRRKKRGLFGAIAGFIEGGW
                                                                 420
QGMVDGWYGYHHSNEQGSGYAADKESTQKAIDGVTNKVNSIIDKMNTQFEAVGREFNNLE
                                                                 480
RRIENLNKKMEDGFLDVWTYNAELLVLMENERTLDFHDSNVKNLYDKVRLQLRDNAKELG
                                                                 540
NGCFEFYHKCDNECMESVRNGTYDYPQYSEEARLKREEISGVKLESIGIYQILSIYSTVA
                                                                 565
SSLALAIMVAGLSLWMCSNGSLQCR (SEQ ID NO: 2)
```

VN04 HA1-1 construct (SEQ ID NO: 11). For this construct four (4) conserved disulfide bonds were preserved by making the amino terminal truncation before the Cysteine at position 58 of SEQ ID NO: 2 and the carboxyl-terminal truncation after the Cysteine at position 318 of SEQ ID NO: 2. The preserved disulfide bonds are listed as the following: C58-C290; C71-C83; C106-C151 and C294-C318 of SEQ ID NO: 2. Using the crystal structure cited above and the Molecular Modeling Database (MMDB) to identify tertiary and secondary structure, the amino terminal truncation was placed at the Glutamic Acid found at position 50 of SEQ ID NO: 2 such that the preceding β-strand [THAQDI, SEQ ID NO: 230, residues 43-48 of SEQ ID NO: 2] is totally eliminated and the subsequent strand [EKKHNG, SEQ ID NO: 231, residues 50-55 of SEQ ID NO: 2] and the β-strand [KLCDLD, SEQ ID NO: 232] formed by residues 56-61 of SEQ ID NO: 2 were left wholly intact. Structurally the random peptide comprising residues 50-55 of SEQ ID NO: 2 [EKKHNG, SEQ ID NO: 231] mimics a β-strand that completes the membrane-distal β-pleated sheet to keep the domain structure intact. The carboxyl-terminal truncation was made in the loop region near residues 323-326 of SEQ ID NO: 2 [KSNR, SEQ ID NO: 233] that follow the Cysteine at position 318 of SEQ ID NO: 2. Truncation within this loop preserves secondary structure in that the preceding β-strand formed by residues 319-322 of SEQ ID NO: 2 [PKYV, SEQ ID

NO: 234] remains wholly intact and the subsequent β-strand [LVLATG, SEQ ID NO: 235, residues 327-332 of SEQ ID NO: 2] is completely eliminated.

VN04 HA1-2 construct (SEQ ID NO: 12). For this construct two (2) conserved disulfide bonds (C71-C83 and C106-C151 of SEQ ID NO: 2) were preserved by making the amino-terminal truncation before the Cysteine at position 71 of SEQ ID NO: 2 and the carboxyl-terminal truncation after the Cysteine at position 151 of SEQ ID NO: 2. Using the crystal structure cited above and the Molecular Modeling Database (MMDB) to identify tertiary and secondary structure, the amino terminal truncation was placed at the glycine at position 62 of SEQ ID NO: 2 such that the subsequent β-strand comprising residues 64-68 of SEQ ID NO: 2 [KPLIL, SEQ ID NO: 236] remains wholly intact, while the preceding β-strand [KLCDLD, SEQ ID NO: 232, residues 56-61 of SEQ ID NO: 2] is completely eliminated.

To preserve secondary structure, the carboxyl-terminal truncation was made in the loop region at the Glutamic Acid at position 284 of SEQ ID NO: 2 such that the preceding β-strand [TIMKS, SEQ ID NO: 237, residues 279-283 of SEQ ID NO: 2] remained wholly or largely intact and the subsequent β-strand [YGNCN, SEQ ID NO: 238, residues 287-291 of SEQ ID NO: 2] was completely eliminated.

VN04 HA1-3 construct (SEQ ID NO: 13). For this construct one (1) conserved disulfide bond (C106-C151 of SEQ ID NO: 2) was preserved by making the amino-terminal truncation before the Cysteine at position 106 of SEQ ID NO: 2 and the carboxyl terminal truncation after the Cysteine at position 151 of SEQ ID NO: 2.

Using the crystal structure cited above and the Molecular Modeling Database (MMDB) to identify tertiary and secondary structure, the amino terminal truncation was placed at the Asparagine at position 103 of SEQ ID NO: 2 such that the preceding β-strand [SYIVEK, SEQ ID NO: 239, residues 93-98 of SEQ ID NO: 2] was completely eliminated while the subsequent β-strand sequence [PGDFN, SEQ ID NO: 240, residues 108-112 of SEQ ID NO: 2] remained wholly or largely intact. To preserve secondary structure, the carboxyl-terminal truncation was made in the random coil region at the Glycine at position 276 of SEQ ID NO: 2 such that the preceding β-strand formed by residues 102-105 of SEQ ID NO: 2 [KGDS, SEQ ID

NO: 241] remains wholly or partially intact, the preceding β-strand [EYAYKIVK, SEQ ID NO: 242, residues 267-274 of SEQ ID NO: 2] is wholly preserved and the subsequent β-strand [TIMKS, SEQ ID NO: 237, residues 279-283 of SEQ ID NO: 2] is completely eliminated.

Depiction of domain boundaries for IND05 HA constructs. The A/Indonesia/5/2005 construct description is based on sequence alignment in reference to A/Viet Nam/1203/2004 structure. The selected boundary domains are highlighted in the sequence below (SEQ ID NO: 3) as follows: HA1-1 boundaries are single-underlined (E50-K323, of SEQ ID NO: 3); HA1-2 boundaries are double-underlined (G62-E284 of SEQ ID NO: 3); HA1-3 boundaries are bold-underlined (N103-G276 of SEQ ID NO: 3). Detailed descriptions of each subunit design and boundary domains are given below.

```
                                                                  60
MEKIVLLLAIVSLVKSDQICIGYHANNSTEQVDTIMEKNVTVTHAQDILEKTHNGKLCDL
                                                                 120
DGVKPLILRDCSVAGWLLGNPMCDEFINVPEWSYIVEKANPTNDLCYPGSFNDYEELKHL
                                                                 180
LSRINHFEKIQIIPKSSWSDHEASSGVSSACPYLGSPSFFRNVVWLIKKNSTYPTIKKSY
                                                                 240
NNTNQEDLLVLWGIHHPNDAAEQTRLYQNPTTYISIGTSTLNQRLVPKIATRSKVNGQSG
                                                                 300
RMEFFWTILKPNDAINFESNGNFIAPEYAYKIVKKGDSAIMKSELEYGNCNTKCQTPMGA
                                                                 360
INSSMPFHNIHPLTIGECPKYVKSNRLVLATGLRNSPQRESRRKKRGLFGAIAGFIEGGW
                                                                 420
QGMVDGWYGYHHSNEQGSGYAADKESTQKAIDGVTNKVNSIIDKMNTQFEAVGREFNNLE
                                                                 480
RRIENLNKKMEDGFLDVWTYNAELLVLMENERTLDFHDSNVKNLYDKVRLQLRDNAKELG
                                                                 540
NGCFEFYHKCDNECMESIRNGTYNYPQYSEEARLKREEISGVKLESIGTYQILSIYSTVA
                                                                 560
SSLALAIMMAGLSLWMCSNGSLQCRICI (SEQ ID NO: 3)
```

IND05 HA1-1 construct (SEQ ID NO: 14). For this construct four (4) conserved disulfide bonds were preserved by making the amino-terminal truncation before the Cysteine at position 58 of SEQ ID NO: 3 and the carboxyl-terminal truncation after the Cysteine at position 318 of SEQ ID NO: 3. The preserved disulfide bonds are listed as the following: C58-C291; C71-C83; C106-C151 and C294-C318 of SEQ ID NO: 3. Using the crystal structure cited above and the Molecular Modeling Database (MMDB) to identify tertiary and secondary structure, the amino terminal truncation was placed at the Glutamic Acid found at position 50

of SEQ ID NO: 3 such that the preceding β-strand [THAQDI, SEQ ID NO: 243, residues 43-48 of SEQ ID NO: 3] was totally eliminated and the subsequent strand [EKTHNG, SEQ ID NO: 244, residues 50-55 of SEQ ID NO: 3] and β-strand [KLCDLD, SEQ ID NO: 245] formed by residues 56-61 of SEQ ID NO: 3 was left wholly intact. Structurally the random peptide comprising residues 50-55 of SEQ ID NO: 3 [EKTHNG, SEQ ID NO: 244] mimics a β-strand that completes the membrane-distal β-pleated sheet to keep the domain structure intact. The carboxyl-terminal truncation was made in the loop region near residues 323-326 of SEQ ID NO: 3 [KSNR, SEQ ID NO: 246] that follow the Cysteine at position 318 of SEQ ID NO: 3. Truncation within this loop preserves secondary structure in that the preceding β-strand formed by residues 319-322 of SEQ ID NO: 3 [PKYV, SEQ ID NO: 247] remains wholly intact and the subsequent β-strand [LVLATG, SEQ ID NO: 248, residues 327-332 of SEQ ID NO: 3] is completely eliminated.

IND05 HA1-2 construct (SEQ ID NO: 15). For this construct two (2) conserved disulfide bonds (C71-C83 and C106-C151 of SEQ ID NO: 3 were preserved by making the amino-terminal truncation before the Cysteine at position 71 of SEQ ID NO: 3 and the carboxyl-terminal truncation after the Cysteine at position 151 of SEQ ID NO: 3. Using the crystal structure cited above and the Molecular Modeling Database (MMDB) to identify tertiary and secondary structure, the amino terminal truncation was placed at the Glycine at position 62 of SEQ ID NO: 3 such that the subsequent β-strand comprising residues 64-68 of SEQ ID NO: 3 [KPLIL, SEQ ID NO: 249] remained wholly intact, while the preceding β-strand [KLCDLD, SEQ ID NO: 245, residues 56-61 of SEQ ID NO: 3] was completely eliminated. To preserve secondary structure, the carboxyl terminal truncation was made in the loop region at the Glutamic Acid at position 284 of SEQ ID NO: 3 in the sequence [ELE, residues 284-286 of SEQ ID NO: 3] such that the preceding β-strand [AIMKS, SEQ ID NO: 250, residues 279-283 of SEQ ID NO: 3] remained wholly or largely intact and the subsequent β-strand [YGNCN, SEQ ID NO: 251, residues 287-291 of SEQ ID NO: 3] was completely eliminated.

IND05 HA1-3 construct (SEQ ID NO: 16). For this construct one (1) conserved disulfide bond (C106-C151 of SEQ ID NO: 3 was preserved by making the amino-terminal truncation before the Cysteine at position 106 of SEQ ID NO: 3

and the carboxyl-terminal truncation after the Cysteine at position 151 of SEQ ID NO: 3. Using the crystal structure cited above and the Molecular Modeling Database (MMDB) to identify tertiary and secondary structure, the amino terminal truncation was placed at the asparagine at position 103 of SEQ ID NO: 3 such that the preceding β-strand [SYIVEK, SEQ ID NO: 252, residues 93-98 of SEQ ID NO: 3] was completely eliminated while the β-strand sequence [PGSFN, SEQ ID NO: 253, residues 108-112 of SEQ ID NO: 3] remained wholly or largely intact. To preserve secondary structure, the carboxyl-terminal truncation was made in the random coil region at the Glycine found at position 276 of SEQ ID NO: 3 such that the preceding β-strand [EYAYKIVK, SEQ ID NO: 254, residues 267-274 of SEQ ID NO: 3] was wholly preserved and the subsequent β-strand [AIMKS, SEQ ID NO: 250, residues 279-283 of SEQ ID NO: 3] was completely eliminated.

Depiction of domain boundaries for New Caledonia HA constructs. A/New Caledonia/20/199 construct description is based on sequence alignment in reference of A/Puerto Rico/8/34 structure. The selected boundary domains are highlighted in the sequence below (SEQ ID NO: 4) as follows: HA1-1 boundaries are single-underlined (S53-R324 of SEQ ID NO: 4); HA1-2 boundaries are double-underlined (K62-S284 of SEQ ID NO: 4); HA1-3 boundaries are bold-underlined (N101-G276 of SEQ ID NO: 4). Detailed descriptions of each subunit design and boundary domains are given below.

```
                                                            60
MKAKLLVLLCTFTATYADTICIGYHANNSTDTVDTVLEKNVTVTHSVNLLEDSHNGKLCL
                                                           120
LKGIAPLQLGNCSVAGWILGNPECELLISKESWSYIVETPNPENGTCYPGYFADYEELRE
                                                           180
QLSSVSSFERFEIFPKESSWPNHTVTGVSASCSHNGKSSFYRNLLWLTGKNGLYPNLSKS
                                                           240
YVNNKEKEVLVLWGVHHPPNIGNQRALYHTENAYVSVVSSHYSRRFTPEIAKRPKVRDQE
                                                           300
GRINYYWTLLEPGDTIIFEANGNLIAPWYAFALSRGFGSGIITSNAPMDECDAKCQTPQG
                                                           360
AINSSLPFQNVHPVTIGECPKYVRSAKLRMVTGLRNIPSIQSRGLFGAIAGFIEGGWTGM
                                                           420
VDGWYGYHHQNEQGSGYAADQKSTQNAINGITNKVNSVIEKMNTQFTAVGKEFNKLERRM
                                                           480
ENLNKKVDDGFLDIWTYNAELLVLLENERTLDFHDSNVKNLYEKVKSQLKNNAKEIGNGC
                                                           540
FEFYHKCNNECMESVKNGTYDYPKYSEESKLNREKIDGVKLESMGVYQILAIYSTVASSL
                                                           565
VLLVSLGAISFWMCSNGSLQCRICI (SEQ ID NO: 4)
```

New Caledonia HA1-1 construct design (SEQ ID NO: 17). For this construct, four (4) conserved disulfide bonds were preserved by making the amino terminal truncation before the Cysteine at position 59 of SEQ ID NO: 4 and the carboxyl-terminal truncation after the Cysteine at position 319 of SEQ ID NO: 4. The preserved disulfide bonds are listed as the following: C59-C291; C72-C84; C107-C152 and C295-C319 of SEQ ID NO: 4. Using the reference crystal structures and the MMDB to check secondary structures the amino terminal truncation was made at the Serine at position 53 of SEQ ID NO: 4 such that the preceding β-strand [THSVNLLED, SEQ ID NO: 255, residues 44-52 of SEQ ID NO: 4] was totally eliminated and the subsequent β-strands SHNGKL, SEQ ID NO: 256 [residues 53-58 of SEQ ID NO: 4] and APLQLG, SEQ ID NO: 257 [residues 65-70, of SEQ ID NO: 4] were left wholly intact. The amino-terminal truncation in the random coil sequence [SHNGKL, SEQ ID NO: 256, residues 53-58 of SEQ ID NO: 4] left the secondary structure largely intact because it mimics a β-strand that leaves the β-pleated sheet intact. The carboxyl-terminal truncation was made at the Arginine at position 324 of SEQ ID NO: 4 in the random coil sequence [RSAK, SEQ ID NO: 258, residues 324-327 of SEQ ID NO: 4] following the Cysteine at position 319 of SEQ ID NO: 4. The truncation in this random coil preserves secondary structure by leaving the preceding β-strand [PKYV, SEQ ID NO: 259, residues 320-323 of SEQ ID NO: 4] wholly intact while the subsequent β-strand [LRMVTG, SEQ ID NO: 260, residues 328-333 of SEQ ID NO: 4] is completely eliminated.

New Caledonia HA1-2 construct design (SEQ ID NO: 18). For this construct two (2) conserved disulfide bonds (C72-C84 and C107-C152 of SEQ ID NO: 4) were preserved by making the amino-terminal truncation before the Cysteine at position 72 of SEQ ID NO: 4 and the carboxyl-terminal truncation after the Cysteine at position 152 of SEQ ID NO: 4. Using the reference crystal structures and the MMDB to check secondary structures the amino-terminal truncation was made at the Lysine at position 62 of SEQ ID NO: 4 such that the random coil sequence comprising [SHNGKLCLLKGI, SEQ ID NO: 261, residues 53-64 of SEQ ID NO: 4] was partially or largely removed, the preceding β-strand [THSVNLLED, SEQ ID NO: 255, residues 44-52 of SEQ ID NO: 4] was completely eliminated and

the subsequent β-strand [APLQLG, SEQ ID NO: 257, residues 65-70 of SEQ ID NO: 4] remained wholly intact. The carboxyl-terminal truncation was made at the Serine at position 284 of SEQ ID NO: 4 such that the preceding β-strand [SGIITS, SEQ ID NO: 262, residues 279-284 of SEQ ID NO: 4] remained wholly intact and the subsequent random coil [NAPMDECDA, SEQ ID NO: 263, residues 285-293 of SEQ ID NO: 4] was wholly or largely eliminated.

New Caledonia HA1-3 construct design (SEQ ID NO: 19). For this construct one (1) conserved disulfide bond (C107-C152 of SEQ ID NO: 4 was preserved by making the amino-terminal truncation before the Cysteine at position 107 of SEQ ID NO: 4 and the carboxyl terminal truncation after C152 of SEQ ID NO: 4. Using the reference crystal structures and the MMDB to check secondary structures the amino-terminal truncation was made at the Asparagine at position 101 of SEQ ID NO: 4 such that the receding β-strand [SYIVET, SEQ ID NO: 264, residues 94-99 of SEQ ID NO: 4] was completely eliminated and the subsequent β-strand [PGYFA, SEQ ID NO: 265, residues 109-113 of SEQ ID NO: 4] remained wholly intact. To preserve the secondary structure the carboxyl-terminal truncation was made at the Glycine at position 276 of SEQ ID NO: 4 such that the β-strand sequence [YAFALSRGF, SEQ ID NO: 266, residues 269-277 of SEQ ID NO: 4] remained largely intact, while the subsequent β-strand sequence [SGIITS, SEQ ID NO: 262, residues 279-284 of SEQ ID NO: 4] was eliminated.

Depiction of domain boundaries for A/Wisconsin/67/2005 HA constructs. A/Wisconsin/67/2005 construct description is based on sequence alignment in reference of A/X31 structure. The selected boundary domains are highlighted in the sequence below (SEQ ID NO: 6) as follows: HA1-1 boundaries are single-underlined (Q44-K310 of SEQ ID NO: 6); HA1-2 boundaries are double-underlined (S54-D271 of SEQ ID NO: 6); HA1-3 boundaries are bold-underlined (S95-G263 of SEQ ID NO: 6). Detailed descriptions of each subunit design and boundary domains are given below.

QKLPGNDNSTATLCLGHHAVPNGTIVKTITNDQIEVTNATELVQSSSTGGICDSPHQILD 60
GENCTLIDALLGDPQCDGFQNKKWDLFVERSKAYSNCYPYDVPDYASLRSLVASSGTLEF 120
NDESFNWTGVTQNGTSSACKRRSNNSFFSRLNWLTHLKFKYPALNVTMPNNEKFDKLYIW 180
GVHHPGTDNDQIFLHAQASGRITVSTKRSQQTVIPNIGSRPRIRNIPSRISIYWTIVKPG 240
DILLINSTGNLIAPRGYFKIRSGKSSIMRSDAPIGKCNSECITPNGSIPNDKPFQNVNRI 300
TYGACPRYVKQNTLKLATGMRNVPEKQTR (SEQ ID NO: 6) 329

A/Wisconsin/67/2005 HA1-1 construct design (SEQ ID NO: 20). For this construct, four (4) conserved disulfide bonds were preserved by making the amino terminal truncation before the Cysteine at position 52 of SEQ ID NO: 6 and the carboxyl-terminal truncation after the Cysteine at position 305 of SEQ ID NO: 6. The preserved disulfide bonds are listed as the following: C52-C277; C64-C76; C97-C139 and C281-C305 of SEQ ID NO: 6. Using the reference crystal structures and the MMDB to check secondary structures the amino terminal truncation was made at the Glutamine at position 44 of SEQ ID NO: 6, right after the tight turn formed by residues 42 and 43 of SEQ ID NO: 6 [LV], such that the preceding β-strand [TNATE, SEQ ID NO: 267, residues 37-41 of SEQ ID NO: 6] was totally eliminated while the entire short β-strand formed by residues QSS [44-46 of SEQ ID NO: 6] was preserved. The carboxyl-terminal truncation was made in the random coil sequence [PRYVK, SEQ ID NO: 268, residues 306-310 of SEQ ID NO: 6] following the Cysteine at position 305 of SEQ ID NO: 6 such that the carboxyl-terminal tail which interacts with HA2 was completely eliminated. Although the carboxyl-terminal sequence [PRYVK, SEQ ID NO:268, residues 306-310 of SEQ ID NO: 6)] is defined as a random coil, structurally it serves as an additional β-strand that further stabilizes the β-pleated sheets defined by residues 44-46 of SEQ ID NO: 6 [QSS] and residues 293-297 of SEQ ID NO: 6 [PFQNV, SEQ ID NO: 269]. This set of β-pleated encompasses the amino-terminus and carboxy-terminus within one stable secondary structure element to ensure a compact domain structure.

A/Wisconsin/67/2005 HA1-2 construct design (SEQ ID NO: 21). For this construct two (2) conserved disulfide bonds (C64-C76 and C97-C139 of SEQ ID NO: 6) were preserved by making the amino-terminal truncation before the Cysteine

at position 64 of SEQ ID NO: 6 and the carboxyl-terminal truncation after the Cysteine at position 139 of SEQ ID NO: 6. Using the reference crystal structures and the MMDB to check secondary structures the amino-terminal truncation was made in the loops that connect two distinct sets of β-pleated sheets. The amino-terminal truncation was made at the Serine at position 54 of SEQ ID NO: 6, such that the β-strand comprising residues 57-62 of SEQ ID NO: 6 [QILDGE, SEQ ID NO: 270] of the membrane-distal set of β-strands remained wholly intact, while the membrane-proximal β-strand [GGICD, SEQ ID NO: 271, residues 49-53 of SEQ ID NO: 6] was completely eliminated. The carboxyl-terminal truncation was made at the Aspartic Acid at position 271 of SEQ ID NO: 6. Truncation at this position preserved secondary structure such that the membrane-distal β-strand [SSIMRS, SEQ ID NO: 272, residues 265-270 of SEQ ID NO: 6] remained intact while the membrane-proximal β-strand [APIGK, SEQ ID NO: 273, residues 272-276 of SEQ ID NO: 6] was eliminated. The membrane-distal β-pleated sheets comprising β-strands QILDGE (SEQ ID NO: 270), [residues 57-62 of SEQ ID NO: 6] LFVER (SEQ ID NO: 274) [residues 86-90 of SEQ ID NO: 6], and SSIMRS, (SEQ ID. NO: 272), [residues 265-270 of SEQ ID NO: 6] serve as the stabilizing secondary structure element that ensures a compact domain structure.

A/Wisconsin/67/2005 HA1-3 construct design (SEQ ID NO: 22). For this construct one (1) conserved disulfide bond (C97-C139 of SEQ ID NO: 6) was preserved by making the amino-terminal truncation before the Cysteine at position 97 of SEQ ID NO: 6 and the carboxy- terminal truncation after the Cysteine at position 139 of SEQ ID NO: 6. Using the reference crystal structure and the MMDB to check secondary structures the amino-terminal truncation was made at the Serine at position 95 of SEQ ID NO: 6 such that the preceding α-helix [TLIDALLG, SEQ ID NO: 275, residues 65-72 of SEQ ID NO: 6] was completely eliminated while the Serine at position 95 of SEQ ID NO: 6 and Asparagine at position 96 of SEQ ID NO: 6 were preserved along with the conserved Cysteine at position 97 of SEQ ID NO: 6. The carboxyl-terminal truncation was made in the random coil sequence at about the Glycine at position 263 of SEQ ID NO: 6 such that the random coil sequence containing residues 261-263 of SEQ ID NO: 6 [RSG] was kept wholly or partially intact, while the preceding β-strand [RGYFKI, SEQ ID NO: 276, residues

255-260 of SEQ ID NO: 6] was wholly preserved and the subsequent β-strand [SSIMRS, SEQ ID NO: 272, residues 265-270 of SEQ ID NO: 6] was completely eliminated.

DISCUSSION

Parts of the globular head domain of the influenza virus hemagglutinin (HA) protein may generate neutralizing antibodies (Brand and Skehel, 1972; Eckert, 1973; Jackson et al., 1979; Russ et al., 1981) and the discrete, globular structure of this domain was thought to constitute a protein fragment which could be expressed in bacteria and refolded, or expressed and secreted from a eukaryotic host, thereby avoiding the expression and purification problems which have been encountered producing the full-length HA protein.

The largest HA construct (also referred to herein as "HA fragment") fused to STF2, HA1-1 (SEQ ID NO: 8, 11, 14, 17, 20), encompasses most of or likely the entire globular region of the naturally occurring (wild type) HA1 head domain. A naturally occurring fragment of HA1 very similar to the HA1-1 construct is released from the virus on limited proteolysis (Bizeband et al., 1995). Thus, structural analysis and experimental data suggest that the HA1-1 construct should fold efficiently, whether produced in bacteria or eukaryotes, and maintain a stable native-like structure. Similarly, it is expected that the entire HA0s, the external portion of HA molecule outside the viral membrane (SEQ ID NO: 23) consisting of HA1 and HA2 but excluding the signal peptide (17 aa), internal peptide within the viral membrane (12 aa) and transmembrane domain (24 aa) should attain its native conformation when expressed in a eukaryotic host or can be refolded into a native state *in vitro* when expressed in bacteria. The two smaller HA fragments (HA1-2 (SEQ ID NO: 9, 12, 15, 18, 21) and HA1-3 (SEQ ID NO: 10, 13, 16, 19, 22)) truncate the region of the head domain distal to the receptor binding site in a manner which may either destabilize the resulting protein fragment (HA1-3) or expose the side chains of a few large hydrophobic residues (HA1-2). For this reason, experimental mutations were introduced into the HA1-2 and HA1-3 constructs, denoted as HA1-2*mut* (SEQ ID NO: 24, 26, 28, 30, 32) and HA1-3*mut* (SEQ ID NO: 25, 27, 29, 31, 33). The substitutions made in the HA1-3*mut* (SEQ ID NO: 25,

27, 29, 31, 33) construct introduce oppositely charged residues on the opposite strands which are left exposed by the HA1-3 truncations. "*HA mut*," as used herein, means an amino acid in the native or naturally occurring HA has been substituted with an amino acid that does not occur in the native or naturally occurring HA. Substituted residues may form a salt bridge and stabilize a structure which may otherwise fold poorly, or not at all. For example, in PR8, G105 was replaced by glutamate and Y115 was substituted with lysine (SEQ ID NO: 25). Negatively charged glutamate interacted with positively charged lysine residue so that N-terminal peptide 1-8 of SEQ ID NO: 25 [NSENEICY, SEQ ID NO: 277] was stabilized by charge-charge interaction with Lysine 15 located at the N-terminus of the short helix 15-22 of SEQ ID NO: 25 [KEELREQL, SEQ ID NO: 278] in the center of the molecule. For the HA1-2 construct, the truncations can lead to the exposure of hydrophobic side chains which would not be expected to impair expression of the protein. However, the exposed residues may lead to aggregation of expressed protein or instability of expressed protein. To avoid aggregation and enhance the stability of the expressed molecule the few large hypdrophobic residues that become exposed were substituted with more neutral amino acids to generate HA1-*2mut* (SEQ ID NO: 24, 26, 28, 30, 32). These substitutions may result in a protein which is more amenable to a robust manufacturing process and/or long term stability.

EXAMPLE 2: DESIGN OF PORTIONS OF A NATURALLY OCCURRING HEMAGGLUTININ OF INFLUENZA B

Materials and Methods

Design of the HA1-1. 1-2 and 1-3 globular head constructs for B/Lee/40.

Structural considerations in the design of an influenza B HA vaccine are similar to those for influenza A, i.e., the domain boundary of the globular head of HA must be identified so that the flagellin-HA fusion protein can fold correctly or be refolded correctly to expose appropriate antigenic epitopes. Unlike influenza A, well-defined X-ray crystallographic structures are not available for influenza B HA, thus it is more difficult to unambiguously define the domain boundary of globular head. Therefore the influenza B HA model must be predicted based on bioinformatic and structural models (Tung et al. 2004. *J. Gen. Virl.* **85**, 3249-3259). These

investigators used a "knowledge-based" approach which depends on a high degree of sequence homology between the known structure from the protein data bank and the target unknown structure. In general, this approach benefits most from at least 35% sequence identity between the known and target proteins.

In the case of influenza B, the closest models come from A/Swine/Hong Kong/9/98 (SEQ ID NO: 34) (24% identity, PDB accession code 1JSD) and A/Aichi/2/68 (SEQ ID NO: 35) (21% identity, PDB accession code 1HGF). Although the sequence identities between the target model B/Lee/40 HA (SEQ ID NO: 36) and known template models are substantially below the desired minimum of 35%, the close similarity of the functions and tertiary folds of influenza A HA proteins in spite of their sequence divergence (H1, H3, H5 and H9 share only 18% sequence identity) suggest a possibly successful prediction of influenza B HA structure using the influenza A HA model. Moreover, influenza C HEF (Hemagglutinin-Esterase-Fusion) protein folds similarly to influenza A HA structure despite even lower sequence identity than any of the A/B comparison.

Since the crystal structure of influenza C HEF is known (PDB accession number 1FLC, MMDB accession number 12663), Tung et al included the knowledge of the structure similarity between C HEF and the known influenza A HA proteins to predict influenza B HA structure. Tung et al first aligned the HEF sequence from C/Johannesburg/1/66 (SEQ ID NO: 37) with one sequence from each of the 15 HA subtypes of influenza A virus (http://flu.lanl.gov) using CLUSTALW (Thompson et al. 1994. *Nucleic Acids Res* **22**, 4673-4680) and compiled a profile based on structure-informed alignment.

The conserved secondary structural features were captured and assigned, as well as the variations among the types and subtypes by year and host species. They then further aligned the augmented A/C profile to an alignment of influenza B HA sequences, including B/Lee/40. The homology modeling techniques of Tung (1999) was used to construct B/Lee/40 model.

Briefly, they first matched the main-chain structures of the target to those of the template in the aligned regions. Insertions in the target relative to the template were treated as loops with known end-structure. Stretching the predicted structure accommodated insertions in the template relative to the target. The mainchain

structures of the loops were modeled by using an efficient Monte Carlo loop-sampling method (Tung, 1997; Ryu et al., 1998). Once the main-chain structure was modeled, the side-chain atoms were attached. As the head of the HA molecule is compact, limited space is available to place the side-chain atoms. Hence, in the analysis, side-chain torsional angles were initialized to equal or be close to those in the template structure. This consideration is particularly useful in avoiding clashes between side chains in the modeled structure. Finally, the all-atom models were subjected to a short run of energy minimization (1000 cycles) by using AMBER (Weiner et al., 1986) to relieve unfavourable steric interactions and to optimize the stereochemistry. The quality of the target model was checked by PROCHECK (Wilson et al.1998. *J. Mol Biol.* **276**, 417-436) and the functionality was checked by substrate docking to test whether the model substrate binding site can accommodate the natural receptor analogue sialyllactose as we have seen in the crystal structure of the HA of A/Aichi/2/68 (Weis et al. 1988. *Nature* **333**, 426-431). The simulated B/Lee/40 model appears reasonable in a Ramachandran Plot (Wilson et al.1998. *J. Mol Biol.* **276**, 417-436) and shows no steric crash when docking the sialyl-2-3-lactose molecule into the substrate binding site of the B/Lee/40 model, indicating a correct stereochemistry. Therefore the B/Lee/40 model was used to guide the design of influenza B HA subunit vaccines.

Results and Discussion

Depiction of Domain Boundaries for B/Lee/40 HA constructs. The selected boundary domains are highlighted in the sequence below as follows: HA1-1 boundaries are single-underlined (T48-K340 of SEQ ID: 36), and HA1-2 boundaries are double-underlined (K60-G299 of SEQ ID: 36). Detailed descriptions of each subunit design and boundary domains are given below.

```
                                                            60
MKAIIVLLMVVTSNADRICTGITSSNSPHVVKTATQGEVNVTGVIPLTTTPTKSHFANLK
                                                           120
GTQTRGKLCPNCFNCTDLDVALGRPKCMGNTPSAKVSILHEVKPATSGCFPIMHDRTKIR
                                                           180
QLPNLLRGYENIRLSTSNVINTETAPGGPYKVGTSGSCPNVANGNGFFNTMAWVIPKDNN
                                                           240
KTAINPVTVEVPYICSEGEDQITVWGFHSDDKTQMERLYGDSNPQKFTSSANGVTTHYVS
                                                           300
QIGGFPNQTEDEGLKQSGRIVVDYMVQKPGKTGTIVYQRGILLPQKVWCASGRSKVIKGS
```

360
LPLIGEADCLHEKYGGLNKSKPYYTGEHAKAIGNCPIWVKTPLKLANGTKYRPPAKLLKE
420
RGFFGAIAGFLEGGWEGMIAGWHGYTSHGAHGVAVAADLKSTQEAINKITKNLNYLSELE
480
VKNLQRLSGAMNELHDEILELDEKVDDLRADTISSQIELAVLLSNEGIINSEDEHLLALE
540
RKLKKMLGPSAVEIGNGCFETKHKCNQTCLDRIAAGTFNAGDFSLPTFDSLNITAASLND
584
DGLDNHTILLYYSTAASSLAVTLMIAIFIVYMVSRDNVSCSICL (SEQ. ID NO: 36)

The coordinates of the simulated B/Lee/40 model are available in Protein Data Bank with the accession code 1TX1. The pdb file was converted to the MMDB (Molecular Modeling Data Base) format by using VAST (Vector Alignment Search Tool) search (http://www.ncbi.nlm.nih.gov/Structure/VAST/vastsearch.html). The model of 1TX1 structure then was viewed by Cn3D and the structure was saved in MMDB format. Based on the same principal used to design influenza A HA1-1 and HA1-2 constructs, the domain boundaries of B/Lee/40 HA-1 and HA 1-2 were pinpointed by examining the model structure. B/Lee/40 HA1-1 (SEQ ID NO: 38) includes the epitope-concentrated globular top, the membrane distal β-sheet and an additional β-sandwich located underneath the membrane distal β-sheet. The amino terminus of HA1-1 starts from residue 48 and ends at residue 340 (TTTPTK (SEQ ID NO: 279)...-...CPIWVK (SEQ ID NO: 280)) of SEQ ID NO: 36). The B/Lee/40 HA1-2 (SEQ ID NO : 39) was designed by removing the β-sandwich from HA1-1. HA1-2 starts from residue 60 and ends at residue 299 (KGTQTR (SEQ ID NO: 281)...-...SKVIKG (SEQ ID NO: 282)) of SEQ ID NO: 36. In order to confirm the boundary selection, other independent methods were also employed. The first method used was the primary sequence alignment. The sequence of A/Aichi/2/68 (SEQ ID NO: 35) was aligned with B/Lee/40 (SEQ ID NO: 36) sequence. The boundaries of A/Aichi/2/68 HA1-1 (SEQ ID NO: 40) (residues 60-326 QSSSTG (SEQ ID NO: 283)...-...CPKYVK (SEQ ID NO: 284)) and HA1-2 (SEQ ID NO: 41) (residues 72-287; HRILDG (SEQ ID NO: 285)...-...SIMRSD (SEQ ID NO: 286)) were aligned closely to those of B/Lee/40, supporting the boundary selections using simulated model. One residue adjustment was made to avoid exposing hydrophobic residues and either end of the construct. Thus, the 3-dimensional structure prediction matches very well with primary sequence alignment.

The second method used was the secondary structure prediction. The domain boundary is usually located in the loop or turn region without invading much of the secondary structure elements such as the α-helix and the β-sheet, especially the center of the α-helix bundle or the β-sheet. This criterion was used to double check if the boundary selection made form simulated 3-D structure is in agreement with independent secondary structure prediction. The program PHD (http://ca.expasy.org/tools → Proteomics and sequence analysis tools → Secondary and tertiary structure tools → PredictProtein) was used to perform the secondary structure prediction. Other than two of the HA1-2 carboxy-terminal residues overlapping with a short α-helix (5 amino acids), all other boundary residues fall in the loop regions, indicating a reasonable boundary selection. PHD results is listed as the following, where AA is the amino acid sequence; OBS_sec is the observed secondary structure: H=helix, E=extended sheet, blank=other (loop); PROF_sec: PROF predicted secondary structure: H=helix, E=extended (sheet), blank=other (loop), PROF = PROF: Profile network prediction HeiDelberg; Rel_sec: reliability index for PROFsec prediction (0=low to 9=high)

For the brief presentation strong predictions marked by '*'; SUB_sec: subset of the PROFsec prediction, for all residues with an expected average accuracy > 82% NOTE: for this subset the following symbols are used: L: is loop (for which above ' ' is used) .: means that no prediction is made for this residue, as the reliability is: Rel < 5; O_3_acc: observerd relative solvent accessibility (acc) in 3 states: b = 0-9%, i = 9-36%, e = 36-100%; P_3_acc: PROF predicted relative solvent accessibility (acc) in 3 states: b = 0-9%, i = 9-36%, e = 36-100%; Rel_acc: reliability index for PROFacc prediction (0=low to 9=high).

For the brief presentation strong predictions marked by '*'; SUB_acc: subset of the PROFacc prediction, for all residues with an expected average correlation > 0.69 (tables in header).

NOTE: for this subset the following symbols are used:

I: is intermediate (for which above ' ' is used)

.: means that no prediction is made for this residue, as the reliability is: Rel < 4.

```
                ....,....1....,....2....,....3....,....4....,....5
AA              MKAIIVLLMVVTSNADRICTGITSSNSPHVVKTATQGEVNVTGVIPLTTT
OBS sec
PROF sec        EEEEEEEEEE     EEEEEEEE     EEEEEE    EEEEEEEEEE
Rel sec         90456542111257762677454037753444441276147765653203
SUB sec         L..EEE......LLLL.EEE.E...LLL.........LL..EEEEEE....

O 3 acc         bbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbb
P 3 acc         ee bbbbbbbbbeeb  bbbbbbb eeee b bb eeebeb  b  bbe
Rel acc         33169997521111032778921201222233322135234021321012
SUB acc         ...bbbbbb........bbbb................e..b.........

                ....,....6....,....7....,....8....,....9....,....1
AA              PTKSHFANLKGTQTRGKLCPNCFNCTDLDVALGRPKCMGNTPSAKVSILH
OBS sec
PROF sec        EEEE                    EEEEEE                EEEE
Rel sec         67752200245344675222033211232342043234337875203787
SUB sec         LLLL......L...LLL.......................LLLL...EEE

O 3 acc         bbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbb
P 3 acc         eee  bbe eeeeeee    eb  bbbbbbbbb b b ee eeeeb bbb
Rel acc         23201011020002300210201031223343310100210221541265
SUB acc         ............................b............eb..bb

                ....,....11.1.,....12.1.,....13.1.,....14.1.,....1
AA              EVKPATSGCFPIMHDRTKIRQLPNLLRGYENIRLSTSNVINTETAPGGPY
OBS sec
PROF sec        EE              HHHHHHHHHHH     EEEEE  EE
Rel sec         50566441224778858887664221002304663021003444556531
SUB sec         E.LLL......LLLLHHHHHHH..........EE.........LLLL..

O 3 acc         bbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbb
P 3 acc         b  eeee bbbbb   e b ebbebbee eeb          eeeeee
Rel acc         12210110221001337264350314531332301200011120132011
SUB acc         ..............e.bi.b...be.....................

                ....,....16.1.,....17.1.,....18.1.,....19.1.,....2
AA              KVGTSGSCPNVANGNGFFNTMAWVIPKDNNKTAINPVTVEVPYICSEGED
OBS sec
PROF sec        HHHHHHHHHHH              EEEEEE
Rel sec         24431001255662036787888752378864236641012244157874
```

```
SUB sec     .........LLLL...HHHHHHHHH..LLLL...LL.........LLLL.

O 3 acc     bbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbb
P 3 acc     bbbbbbee ee bbbbbbb  bbeeeee     ee ebeb b b eee
Rel acc     11111012110301026314921522122321112112322203203212
SUB acc     ................b..bb..b.........................

            ....,....21.1.,....22.1.,....23.1.,....24.1.,....2
AA          QITVWGFHSDDKTQMERLYGDSNPQKFTSSANGVTTHYVSQIGGFPNQTE
OBS sec
PROF sec    EEEEEEEE      HHHHHH       EEEEE  EEEEEEEE
Rel sec     47897532588742544320368853775200122343201275331035
SUB sec     .EEEEE..LLLL..H......LLLL.EEE............LL.....L

O 3 acc     bbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbb
P 3 acc     bbbbbbbbb eeee     b eeeeeeb b b e e e    ebee b    e
Rel acc     19696993033223122210203213224011122211311001202001
SUB acc     .bbbbbb.....................b.....................

            ....,....26.1.,....27.1.,....28.1.,....29.1.,....3
AA          DEGLKQSGRIVVDYMVQKPGKTGTIVYQRGILLPQKVWCASGRSKVIKGS
OBS sec
PROF sec    EEEEEEEEE     EEEEEE    EEE  EEEEE     EEEE
Rel sec     56754320377888983577247886434023022235520330245401
SUB sec     LLLL.....EEEEEEE.LLL..EEEE...........EE.......E...

O 3 acc     bbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbb
P 3 acc     eee  bbbbbbb bbbb ee  b b   bbbbbb  bb b eeb bbe e
Rel acc     17211105270744553301222473010244520230221020211103
SUB acc     .e.....b.b.bibbb.......ib.....bbb................

            ....,....31.1.,....32.1.,....33.1.,....34.1.,....3
AA          LPLIGEADCLHEKYGGLNKSKPYYTGEHAKAIGNCPIWVKTPLKLANGTK
OBS sec
PROF sec    HHH     HHHHH HHHHH
Rel sec     32233333000113432246764334340342334133220003551573
SUB sec     ..................LLL.....................HH.LL.

O 3 acc     bbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbb
P 3 acc     e     e    e bb   ee  b      bbbebbe be bb bbbbb
Rel acc     10010011112110110111101210102102111021712452461321
SUB acc     ....................................b..bb.bb....
```

```
                ....,....36.1.,....37.1.,....38.1.,....39.1.,....4
AA              YRPPAKLLKERGFFGAIAGFLEGGWEGMIAGWHGYTSHGAHGVAVAADLK
OBS sec
PROF sec        EEEEHHHH                          EEEEE  H
Rel sec         56853211324203110113404531100120021023564325530623
SUB sec         LLLL.....................L.............LL...EE..L..

O 3 acc         bbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbb
P 3 acc         b  eee be  bbbbbbbbbbbbb bbbbbbbbbbbbb bbbbbbb e
Rel acc         00011411202355999999623520456353043010011315264032
SUB acc         .....e......bbbbbbbbb..b..bbb.b..b.........b.bb...

                ....,....41.1.,....42.1.,....43.1.,....44.1.,....4
AA              STQEAINKITKNLNYLSELEVKNLQRLSGAMNELHDEILELDEKVDDLRA
OBS sec
PROF sec        HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH HHH
Rel sec         47888989873212477733101234446677898888878756210111
SUB sec         .HHHHHHHHH.....HHH..........HHHHHHHHHHHHHHHH......

O 3 acc         bbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbb
P 3 acc         bbbebbeebbe b  bbee beebeeb ebbeebbe beeb eebee
Rel acc         28049825725342322312063124534063444467553254223101
SUB acc         .b.ebb.eb.e.b........e...eb.e.b.ebbeibee..ee......

                ....,....46.1.,....47.1.,....48.1.,....49.1.,....5
AA              DTISSQIELAVLLSNEGIINSEDEHLLALERKLKKMLGPSAVEIGNGCFE
OBS sec
PROF sec        HHHHHHHHHHHHHH          HHHHHHHHHHHHHHHHHHHH    EEE
Rel sec         03645477777753464114550578879888887664424540663252
SUB sec         ..H.H.HHHHHHH..L....LL.HHHHHHHHHHHHHH....H..LL..E.

O 3 acc         bbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbb
P 3 acc         bbbb bb bbbbb  eebb  e e b eb e b e b e bee eebbbb
Rel acc         04450592968651122001021323439443944173428232012340
SUB acc         .bbb.bb.bbbbb............i.bie.bie.b.e.b.......b.
```

```
                ....,....51.1.,....52.1.,....53.1.,....54.1.,....5
                TKHKCNQTCLDRIAAGTFNAGDFSLPTFDSLNITAASLNDDGLDNHTILL
AA              (SEQ ID NO: 36)
OBS_sec
PROF_sec        EEEE  HHHHHHHH           HHHH  EEEEEEEE      EEEEE
Rel_sec         22003402344311355466523223103101002476217776303665
SUB_sec         ..............LL.LLL.............EE..LLLL...EEE

O_3_acc         bbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbb
P_3_acc         b beb e bbbbb bbbb  ee  eebbe beb bbeb eeeeeeb bbb
Rel_acc         42120312620271012211330320211224131235303222133868
SUB_acc         b.......b...b.................e.....b.........bbb

                ....,....56.1.,....57.1.,....58.1.   (SEQ ID NO: 36)
AA              YYSTAASSLAVTLMIAIFIVYMVSRDNVSCSICL
OBS_sec
PROF_sec        EEEHHHHHHHHHHHHHHHHHH      EEEEEEE
Rel_sec         43100214788888887763210067403468 60
SUB_sec         ........HHHHHHHHHHH.....LL....EEE.

O_3_acc         bbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbb
P_3_acc         bbbbbbbbbbbbbbbbbbbbbbee eb b bbb
Rel_acc         63623666799999969985171010101524 00
SUB_acc         b.b..bbbbbbbbbbbbbb.b.......b.b..
```

The third tool that used is the hydrophobicity analysis. In general, the structural core of a protein tends to be a hydrophobic cluster flanked by stretch of hydrophilic residues. Therefore the predicted boundaries should not locate in the center of hydrophobic core, rather should reside in the flanking hydrophilic regions. Hydrophobicity of B/Lee/40 HA sequence (SEQ ID NO: 36) was plotted using ProtScale (http://ca.expasy.org/tools → Primary structure analysis → ProtScale). The hydrophobicity analysis of B/Lee/40 sequence confirmed that the selected boundaries were in the hydrophilic regions of the protein (Figure 2). The hydrophobicity strength used in the plot is listed in the following table:

SEQUENCE LENGTH: 584

Using the scale **Hphob. / Kyte & Doolittle**, the individual values for the 20 amino acids are:

Ala:	1.800	Arg:	-4.500	Asn:	-3.500	Asp:	-3.500	Cys:	2.500	Gln:	-3.500
Glu:	-3.500	Gly:	-0.400	His:	-3.200	Ile:	4.500	Leu:	3.800	Lys:	-3.900
Met:	1.900	Phe:	2.800	Pro:	-1.600	Ser:	-0.800	Thr:	-0.700	Trp:	-0.900
Tyr:	-1.300	Val:	4.200	Asx:	-3.500	Glx:	-3.500	Xaa:	-0.490		

In order to identify the HA 1-1 and HA1-2 domain boundaries of other influenza B molecules, such as B/Malaysia/2506/2004 (SEQ ID NO: 42), B/Ohio/1/2005 (SEQ ID NO: 43), B/Victoria/2/87 lineage (SEQ ID NO: 787) and B/Shanghai/361/2002 (SEQ ID NO: 44), and B/Yamagata/16/88 lineage (SEQ ID NO: 213), each HA sequence was aligned with B/Lee/40 HA sequence (SEQ ID NO: 36) . The results indicated that influenza B HA proteins are much more conserved compared to influenza A HA proteins, especially in the domain boundary sequences. Thus, the boundaries identified in B/Lee/40 were used to direct boundary selections of other influenza B strains. The resulting domain boundaries for each strain are summarized below.

B/Malaysia/2506/2004: HA1-1 44-337 (SEQ ID NO: 45) (TTTPTK (SEQ ID NO: 287)-CPIWVK (SEQ ID NO: 288)) HA1-2 56-296 (SEQ ID NO: 46) (KGTETR (SEQ ID NO: 289)-SKVIKG (SEQ ID NO: 290))

B/Ohio/1/2005: HA1-1 33-326 (SEQ ID NO: 47) (TTTPTK (SEQ ID NO: 291)-CPIWVK (SEQ ID NO: 292)) HA1-2 45-285 (SEQ ID NO: 48) (KGTKTR (SEQ ID NO: 293)-SKVIKG (SEQ ID NO: 294))

B/Shanghai/361/2002: HA1-1 33-325 (SEQ ID NO: 49) (TTTPIK (SEQ ID NO: 295)-CPIWVK (SEQ ID NO: 296)) HA1-2 45-284 (SEQ ID NO: 50) (KGTRTR (SEQ ID NO: 297)-SKVIKG (SEQ ID NO: 298))

```
                                                                    50
BMal        ----IVLLMVVTSNADRICTGITSSNSPHVVKTATQGEVNVTGVIPLTTT
BOhi        ---------------DRICTGITSSNSPHVVKTATQGEVNVTGVIPLTTT
BLee        MKAIIVLLMVVTSNADRICTGITSSNSPHVVKTATQGEVNVTGVIPLTTT
BSha        ---------------DRICTGITSSNSPHVVKTATQGEVNVTGVIPLTTT
                           ************************************
```

```
                                                                      100
     BMal       PTKSHFANLKGTETRGKLCPKCLNCTDLDVALGRPKCTGNIPSARVSILH
     BOhi       PTKSHFANLKGTKTRGKLCPKCLNCTDLDVALGRPKCTGNIPSAEVSILH
     BLee       PTKSHFANLKGTQTRGKLCPNCFNCTDLDVALGRPKCMGNTPSAKVSILH
     BSha       PIKSHFANLKGTRTRGKLCPDCLNCTDLDVALGRPMCVGTTPSAKASILH
                * **********.*******.*:************ * *. ***..****

                                                                      150
     BMal       EVRPVTSGCFPIMHDRTKIRQLPNLLRGYEHIRLSTHNVINAENAPGGSY
     BOhi       EVRPVTSGCFPIMHDRTKIRQLPNLLRGYEHIRLSTHNVINAEKAPGGPY
     BLee       EVKPATSGCFPIMHDRTKIRQLPNLLRGYENIRLSTSNVINTETAPGGPY
     BSha       EVRPVTSGCFPIMHDRTKIRQLPNLLRGYENIRLSTQNVIDAEKALGGPY
                **:*.****************************:***** ***::*.* **.*

                                                                      200
     BMal       KIGTSGSCPNVTNGNGFFATMAWAVPKNDNNKTATNSLTIEVPYICTEGE
     BOhi       KIGTSGSCPNVTNGNGFFATMAWAVPKNDNNKTATNSLTIEVPYICTEGE
     BLee       KVGTSGSCPNVANGNGFFNTMAWVIPK-DNNKTAINPVTVEVPYICSEGE
     BSha       RLGTSGSCPNATSKSGFFATMAWAVPK-DNNKNATNPLTVEVPYICTEGE
                ::********.:. .*** ****.:** ****.* *.:*:*******:***

                                                                      250
     BMal       DQITVWGFHSDNEAQMAKLYGDSKPQKFTSSANGVTTHYVSQIGGFPNQT
     BOhi       DQITIWGFHSDSETQMAKLYGDSKPQKFTSSANGVTTHYVSQIGGFPNQT
     BLee       DQITVWGFHSDDKTQMERLYGDSNPQKFTSSANGVTTHYVSQIGGFPNQT
     BSha       DQITVWGFHSDDKTQMKNLYGDSNPQKFTSSANGVTTHYVSQIGGFPDQT
                ****:*******.::** .******:************************:**

                                                                      300
     BMal       EDGGLPQSGRIVVDYMVQKSGKTGTITYQRGILLPQKVWCASGRSKVIKG
     BOhi       EDGGLPQSGRIVVDYMVQKSGKTGTITYQRGILLPQKVWCASGRSKVIKG
     BLee       EDEGLKQSGRIVVDYMVQKPGKTGTIVYQRGILLPQKVWCASGRSKVIKG
     BSha       EDGGLPQSGRIVVDYMVQKPGKTGTIVYQRGVLLPQKVWCASGRSKVIKG
                ** ** *************.******.****:******************

                                                                      350
     BMal       SLPLIGEADCLHEKYGGLNKSKPYYTGEHAKAIGNCPIWVKTPLKLANGT
     BOhi       SLPLIGEADCLHEKYGGLNKSKPYYTGEHAKAIGNCPIWVKTPLKLANGT
     BLee       SLPLIGEADCLHEKYGGLNKSKPYYTGEHAKAIGNCPIWVKTPLKLANGT
     BSha       SLPLIGEADCLHEKYGGLNKSKPYYTGEHAKAIGHCPIWVKTPLKLANGT
                ************************************:***************

                                                                      400
     BMal       KYRPPAKLLKER----------------------------------------
     (SEQ ID NO: 42)
     BOhi       KYRPPAKLLKERGF--------------------------------------
     (SEQ ID NO: 43)
     BLee       KYRPPAKLLKERGFFGAIAGFLEGGWEGMIAGWHGYTSHGAHGVAVAADL
     (SEQ ID NO: 36)
     BSha       KYRP------------------------------------------------
     (SEQ ID NO: 44)
                ****
```

B/Lee/40 HA1-1 construct design (SEQ ID NO: 38). For this construct, five (5) anticipated conserved disulfide bonds were preserved by making the amino terminal truncation before the Cysteine at position 69 and the carboxyl-terminal

truncation after the Cysteine at position 335 of SEQ ID NO: 36. The likely disulfide pairings are as the following: C69 and C72, C75 and C87, C109 and C158, C195 and C289, C309 and C335 of SEQ ID NO: 36. Using the simulated structures and the MMDB to check secondary structures the amino terminal truncation was made at the Theonine at position 48, right after the β-strand kink introduced by Proline 46 followed by Leucine 47 of SEQ ID NO: 36, such that the preceding β-strand [TGVI, SEQ ID NO: 299, residues 42-45 of SEQ ID NO: 36] was totally eliminated while the entire short β-strand formed by residues TTTP (SEQ ID NO: 300) [residues 48-51 of SEQ ID NO: 36] was preserved. The carboxyl-terminal truncation was made at the end of β-strand PIWV (SEQ ID NO: 301) [residues 336-340 of SEQ ID NO: 36] following the Cysteine at position 335 such that the carboxyl-terminal tail which interacts with HA2 was completely eliminated. Six β-strands including the C-terminal β-strand form a stable β-sandwich defined by β-strands PIWV (SEQ ID NO: 301) [residue 336-340 of SEQ ID NO: 36], PYYTG (SEQ ID NO: 302) [residues 322 to 326 of SEQ ID NO: 36] and TTTP (SEQ ID NO: 300) [residues 48-51 of SEQ ID NO: 36] (lower β-pleated sheet), and the β-strands DCLHE (SEQ ID NO: 303) [residues 308-312 of SEQ ID NO: 38], YGGLN (SEQ ID NO: 304) [residue 314-318 of SEQ ID NO: 36] and KAIGN (SEQ ID NO: 305) [residues 330-334 of SEQ ID NO: 36] (top β-pleated sheet). This set of β-pleated sheet includes both the amino-terminus and carboxyl-terminus within one stable secondary structure element. Such truncations were made to believe resulting a compact domain structure.

B/Lee/40 HA1-2 construct design (SEQ ID NO: 39). For this construct four (4) anticipated conserved disulfide bonds were preserved by making the amino-terminal truncation before the Cysteine at position 69 and the carboxyl-terminal truncation after the Cysteine at position 289 of SEQ ID NO: 36. Using the simulated structure and the MMDB to check secondary structures the amino-terminal truncation was made in the loops that connect two distinct sets of β-pleated sheets. The amino-terminal truncation was made at the Lysine at position 60, such that the β-strand comprising residues 62-66 of SEQ ID NO: 36 [TQTRG, SEQ ID NO: 306] of the membrane-distal set of β-strands remained wholly intact, while the membrane-proximal β-strand [TTTP, SEQ ID NO: 300, residues 48-51 of

SEQ ID NO: 36] that forms part of HA1-1 was completely eliminated. The carboxyl-terminal truncation was made at the Serine 300 of SEQ ID NO: 36. Truncation at this position preserved secondary structure such that the membrane-distal β-strand [KVIKG, SEQ ID NO: 307, residues 295-299 of SEQ ID NO: 36] remained intact while the membrane-proximal β-strand [DCLHE, SEQ ID NO: 308, residues 308-312 of SEQ ID NO: 36] was eliminated. The membrane-distal β-pleated sheets comprising β-strands TQTRG (SEQ ID NO: 306) [residues 62-66 of SEQ ID NO: 36], SILHEV (SEQ ID NO: 309) [residues 97-102 of SEQ ID NO: 36], and KVIKG (SEQ ID NO: 307) [residues 295-299 of SEQ ID NO: 36] serve as the stabilizing secondary structure element that is believed to conclude a compact domain structure.

B/Malaysia/2506/2004, B/Ohio/1/2005 and B/Shanghai/361/2002 follow the same reference structure as described above.

EXAMPLE 3: CLONING AND EXPRESSION OF RECOMBINANT FLAGELLIN-HEMAGGLUTININ FUSION PROTEINS IN *E. coli*

Materials and Methods

Cloning of HA influenza A subunits. Subunits of the HA globular head from several strains of influenza A were cloned and expressed alone or as fusions with flagellin. We also expressed in fusion with the HA globular head domain, two proteins widely used as carrier proteins in conjugated vaccines. CRM-197 is a mutated diphtheria toxin (DTx) from *Corynebacterium diphtheriae* and LTB, is *E. coli* heat labile toxin B subunit.

These constructs were generated in one of four different methodologies:

Method #1: In this protocol, a fusion gene comprising flagellin (STF2) (SEQ ID NO: 212) and the HA subunit was codon-optimized for *E. coli* expression and obtained from a commercial vendor (DNA2.0 Inc., Menlo Park, CA) by chemical synthesis. The gene was excised with *Nde*I and *Blp*I enzymes, the insert fragment was gel purified and ligated to pET24a (Novagen, San Diego, CA) which had been digested with *Nde*I and *Blp*I and treated with bacterial alkaline phosphatase (BAP).

Method #2: For facile cloning of genes in fusion with flagellin, a cassette plasmid containing a unique *Blp*I site at the 3' end of the flagellin (STF2) gene (SEQ

ID NO: 212) was generated. This was done by introducing a silent mutation at nucleotides 5'-GTGCTGAGCCTGTTACGT - 3' (SEQ ID NO: 310) [nt 1501 to 1518 of SEQ ID NO: 212] of STF2 creating the unique *Blp*I site in the plasmid cassette, pET24/STF2.blp (SEQ ID NO: 51). Synthetic genes for each target antigen were codon-optimized for *E. coli* expression and obtained from a commercial vendor (DNA2.0 Inc., Menlo Park, CA). The synthetic genes were excised with *Blp*I enzyme and ligated by compatible ends to pET24/STF2.blp which had been treated with *Blp*I and BAP.

Method #3: Using forward and reverse primers (Keck Foundation BRLK, Yale University, New Haven, CT; Midland Certified Reagent Company, Midland, TX) as indicated for each construct, PCR amplification was performed using the DNA templates shown in each table. The PCR product was subjected to *Blp*I digestion, gel-purified and ligated to pET24/STF2.blp (SEQ ID NO: 51) vector previously prepared by *Blp*I digestion and BAP treatment.

Method #4: A plasmid cassette, pET24/STF2.SG (SEQ ID NO: 52), was generated by introducing a flexible heptamer linker, Ser-Gly-Ser-Gly-Ser-Gly-Ser (S-G-S-G-S-G-S (SEQ ID NO: 311)) at the 3'end of the STF2 gene. A unique *BamH*I site was created in the linker to facilitate cloning of HA subunit fragments in fusion with flagellin. Synthetic genes codon-optimized for expression in *E. coli* were obtained from a commercial vendor (DNA2.0 Inc., Menlo Park, CA), excised from the parental plasmids with *BamH*I and *Blp*I restriction endonucleases, and ligated by compatible ends to the pET24/STF2 cassette that had previously been digested with *Bam*HI and BAP-treated.

In each case, the constructed plasmids were used to transform competent *E. coli* TOP10 cells and putative recombinants were identified by PCR screening and restriction mapping analysis. The integrity of the constructs was verified by DNA sequencing and they were used to transform the expression host, BLR3 (DE3) (Novagen, San Diego, CA; Cat #69053). Transformants were selected on plates containing kanamycin (50 μg/mL), tetracycline (5 μg/mL) and glucose (0.5%). Colonies were picked and inoculated into 2 ml of LB medium supplemented with 25 μg/ml kanamycin, 12.5 μg/ml tetracycline and 0.5% glucose and grown overnight. Aliquots of these cultures were used to inoculate fresh cultures in the same medium

formulation, which were cultured until they reached an OD_{600} = 0.6, at which time protein expression was induced by the addition of 1 mM IPTG and culturing for 3 hours at 37°C. The cells were harvested and analyzed for protein expression.

SDS-PAGE and Western blot: Protein expression and identity were determined by gel electrophoresis and immunoblot analysis. Cells were harvested by centrifugation and lysed in Laemmli buffer. An aliquot of 10 µl of each lysate was diluted in SDS-PAGE sample buffer with or without 100 mM DTT as a reductant. The samples were boiled for 5 minutes and loaded onto a 10% SDS polyacrylamide gel and electrophoresed (SDS-PAGE). The gel was stained with Coomassie R-250 (Bio-Rad; Hercules, CA) to visualize protein bands. For western blot, 0.5 µl/lane cell lysate was electrophoresed and electrotransfered to a PVDF membrane and blocked with 5% (w/v) dry milk.

The membrane was then probed with either anti-flagellin antibody (Inotek; Beverly, MA) or influenza A PR/8/38 convalescent immune mouse serum. PR/8/34 immune serum was generated in BALB/c mice (Jackson Laboratory, Bar Harbor, ME) that received an experimentally determined sublethal challenge dose of $8x10^1$ egg infectious dosages (EID) of PR/8/34 influenza virus. Animals were then allowed to convalesce for >21days post-infection at which time immune serum was isolated and clarified. After probing with alkaline phosophatase-conjugated secondary antibodies (Pierce; Rockland, IL), protein bands were visualized with an alkaline phosphatase chromogenic substrate (Promega, Madison, WI). Bacterial clones which yielded protein bands of the correct molecular weight and reactive with the appropriate antibodies were selected for production of protein for use in biological assays.

The constructs derived from the HA of A/Puerto Rico/8/34 strain (PR8) (SEQ ID NO: 1) and listed in Table 1 were made by the synthetic gene route as described in Method #1 and Method #4. Similarly, the constructs derived from the A/Viet Nam/1203/2004 strain (SEQ ID NO: 2) are shown in Table 2, the constructs derived from the HA of A/Indonesia/2005 strain (IND) (SEQ ID NO: 3) are described in Table 3, and those from A/New Caledonia/12/99 strain (NC) (SEQ ID NO: 4) are described in Tables 4. "IND" as used herein, refers to "Indonesia."

Where appropriate, the DNA primers and DNA templates used in the PCR amplification reaction are listed in the same table.

Table 1: PR8 HA constructs for expression in *E coli*

SEQ ID NO:	Construct	Method
53	STF2.HA1-1	#4
54	STF2.HA1-1.his	#4
55	STF2.HA1-2	#1
56	STF2.HA1-2*mut*	#1
57	STF2.HA1-3	#1
58	STF2.HA1-3*mut*	#1
59	HA1-1	#1
60	HA1-1.his	#1
61	HA1-2.his	#1
62	CRM.HA1-2	#1
63	LTB.HA1-2	#1

Table 2: VN HA constructs for expression in *E coli*

SEQ ID NO:	Construct	Method	FOR Primer SEQ ID NO:	REV Primer SEQ ID NO:	DNA Templat SEQ ID NO:
64	STF2.HA1-1	#3	65	66	67
68	STF2.HA1-2	#3	69	70	67
71	STF2.HA1-2*mut*	#2	N/A	N/A	72

Table 3: IND HA constructs for expression in *E coli*

SEQ ID NO:	Construct	Method	FOR Primer SEQ ID NO:	REV Primer SEQ ID NO:	DNA Templat SEQ ID NO:
73	STF2.HA1-1	#3	74	66	75
76	STF2.HA1-2	#3	69	77	75
78	STF2.HA1-2*mut*	#2	N/A	N/A	79

Table 4: NC HA constructs for expression in *E coli*

SEQ ID NO:	Construct	Method	FOR Primer SEQ ID NO:	REV Primer SEQ ID NO:	DNA Templat SEQ ID NO:
80	STF2.HA1-1	#3	81	82	83
84	STF2.HA1-2	#3	85	86	83
87	STF2.HA1-2*mut*	#2	N/A	N/A	88

Results

Protein carriers have widespread application in human vaccines. The cross-reactive material (CRM_{197}) of diphtheria toxin is considered to be advantageous as a carrier molecule in the formulation of several conjugate vaccines. Exemplary carriers include *E. coli* heat labile enterotoxin (LT) and its B subunit (LTB), Tetanus toxoid (TT) and cholera toxin (CT). Using CRM_{197} and LTB as representatives of this group of carrier proteins, we have generated constructs in which the globular head of *HA* of (A/Puerto Rico/8/34 strain (PR8) (SEQ ID NO: 1) is fused to the 3'end of either CRM197 gene or LTB gene as described in Method #1 generating the constructs CRM.HA1-2 (SEQ ID NO: 62) and LTB.HA1-2(SEQ ID NO: 63). The constructs were verified by DNA sequencing and used to transform the expression host, BLR3 (DE3) (Novagen, San Diego, CA; Cat #69053).

Transformants were selected on plates containing kanamycin (50 µg/mL), tetracycline (5 µg/mL) and glucose (0.5%). Several colonies were picked for an overnight culture which were was used to inoculate a fresh LB medium supplemented with 25 µg/ml kanamycin, 12.5 µg/ml tetracycline and 0.5% glucose. At an OD600 = 0.6 protein expression was induced with 1 mM IPTG for 3 h at 37°C. The cells were harvested and an aliquot of the lysate was analyzed on 10% SDS-PAGE by Coomassie blue staining and by immunoblot using PR/8/34 convalescent sera. In the case of CRM.HA1-2 construct several clones were picked and analyzed for expression by SDS-PAGE.

As assayed by Coomassie blue staining of the SDS-PAGE gel, all the clones displayed a band that migrated with an apparent MW 84 KDa and that corresponds to the predicted MW. The absence of this band in the control culture (without IPTG) indicates that it is specifically induced by IPTG. This observation was further

confirmed when the cell extracts of two clones #5 and #6 was fractionated in the presence or absence of a reducing agent (5 mM DTT). While the recombinant protein whose disulphide bonds have been disrupted by treatment with DTT is not recognized by the cognate antibodies, the native recombinant protein is. Similarly, a clone expressing construct LTB.HA1-2 displays a band corresponding to the predicted molecular weight of 39.8 KDa, protein when induced with IPTG. The identity of LTB-HA1-2 fusion protein is confirmed by western blot analysis using mouse convalescent serum. This band is diminished in intensity when a reductant (β-mercaptoethanol) was present. This latter observation suggests that insufficient amount of reductant was most likely employed in the experiment. Taken together the data presented herein support the notion that the globular head of HA can be successfully fused to carrier proteins to generate conformational sensitive proteins.

CLONING OF RECOMBINANT FLAGELLIN-HEMAGGLUTININ FUSION PROTEINS IN *E. coli*

Cloning of HA Influenza B subunits. Subunits of the HA globular head from several strains of influenza B were cloned and expressed as fusions with flagellin. These constructs were generated by a two-step PCR.

The HA subunit was codon-optimized for *E. coli* expression and obtained from a commercial vendor (DNA2.0 Inc., Menlo Park, CA) by chemical synthesis. The HA1-1 or HA1-2 was PCR amplified using the synthesized DNA as templates. The flagellin (STF2) sequence (SEQ ID NO: 212) was derived from the plasmid pET24a-STF2.HA1-2. The STF2 DNA fragment was PCR amplified, and the C-terminal of the PCR product has a 28-30 bp overlap with the N-terminal sequence of the fusion HA subunit.

The STF2 and HA subunits were fused together by a 2nd PCR. Using respective forward and reverse primers (Integrated DNA Technologies, Inc, Coralville, IA 52241) listed below, fusion protein DNA was amplified from the DNA templates also shown below. The PCR product was subsequently subjected to *Xba*I digestion, gel-purification and ligation to pET24a-STF2.HA1-2 that was previously digested with XbaI and SnaBI.

The constructs are listed below, where appropriate, the DNA primers and DNA templates used in the PCR amplification reaction are also listed.

FluB STF2.HA constructs for expression in *E coli*

SEQ ID NO:	Construct	FOR Primer SEQ ID NO:	REV Primer SEQ ID NO:	DNA Template SEQ ID NO:
184	STF2.HA1-1 (MAL)	193, 195	194, 196	55, 183
186	STF2.HA1-2 (MAL)	193, 198	197, 199	55, 185
188	STF2.HA1-2 (SH)	193, 201	200, 202	55, 187
190	STF2.HA1-2 (Lee)	193, 204	203, 205	55, 189
192	STF2.HA1-2 (Ohio)	193, 207	206, 208	55, 191

In each case, the constructed plasmids were used to transform competent *E. coli* DH5α cells and putative recombinants were identified by restriction mapping. The integrity of the constructs was verified by DNA sequencing and was used to transform the expression host, BLR3 (DE3) (Novagen, San Diego, CA; Cat #69053). Transformants were selected on LB (Veggie peptone, EMD Bioscience, La Jolla, CA; Cat#71280; veggie yeast extract, EMD Bioscience, Cat# 71279;) plates containing kanamycin (35 μg/mL). Colonies were picked and inoculated into 2 ml of LB (Veggie) medium supplemented with 35 μg/ml kanamycin and grown overnight. Aliquots of these cultures were used to inoculate fresh cultures in the same medium formulation, which were cultured until OD_{600} = 0.6 was reached. The protein expression was induced by the addition of 1 mM IPTG. The cells were harvested after 2 hours culturing at 37°C and analyzed for protein expression.

SDS-PAGE and Western blot: Protein expression and identity were determined by gel electrophoresis and immunoblot analysis. Cells were harvested by centrifugation and lysed in Tris/NaCl (50mM Tris, 200 mM NaCl, pH 8.0) buffer. Two aliquots of each lysate was diluted in SDS-PAGE sample buffer and boiled for 5 minutes. One aliquot sample was reduced by DTT (200mM final concentration, EMD Bioscience, La Jolla, CA 92039; Cat#233155). The samples (1.5 μl/lane cell lysate) were loaded onto a 4-12 % SDS polyacrylamide gel and electrophoresed (SDS-PAGE). The gel was stained with Coomassie R-250 (Bio-Rad; Hercules, CA) to visualize protein bands.

For western blot, the duplicated gel was electrophoresed and electrotransfered to a PVDF membrane and blocked with 5 % (w/v) dry milk. The

membrane was then probed with either anti-flagellin antibody (6H11, Inotek, Beverly, MA; Lot#1030B5) or ferret antiserum raised against B/Malaysia (2506/2004, CDC#2005741987). After probing with Horseradish Peroxidase (HRP)-conjugated secondary antibodies (1:10,000) (Goat anti-Mouse is from Jackson ImmunoResearch Laboratories; West Grove, PA. Goat anti-Ferret is from Bethyl Laboratories; Montgomery, TX.), protein bands were visualized after with addition of TMB, an HRP chromogenic substrate (Pierce; Rockland, IL; Cat#34018). The approximate molecular weight of STF2.HA1-1(MAL) is 84 kDa; and the approximate molecular weight of STF2.HA1-2 (MAL and SH) is ~78 kDa. Two micrograms of STF2.4×M2e fusion protein was loaded as a positive control for the anti-flagellin antibody and a negative control for the anti-HA sera. Virus lysate (B/Malaysia/2506/04, CDC# 2005756250) was included in the experiment as a positive control for the anti-HA sera and a negative control for the anti-flagellin antibody. Commassie blue staining of the gel showed low induction of STF2.HA1-1(MAL) protein; but much stronger bands for STF2.HA1-2(MAL) and STF2.HA1-2(SH) proteins. Western blot analysis using anti-flagellin monoclonal antibodies (1:4000) exhibited a strong positive band around 84 kDa for lanes loaded with STF2.HA1-1(MAL), and around 75 kDa in lanes loaded with STF2.HA1-2(MAL) or STF2.HA1-2(SH). The induction pattern of each fusion protein on the Western is very similar to the result for the commassie blue stained gel. The level of the non-reduced and the reduced protein recognized by the anti-flagellin was comparable. The positive control for anti-flagellin (STF2.4×M2e) showed s single strong band, and the negative control exhibited no signal.

Western blot analysis using anti-HA polyclonal anti-sera (1:1000) exhibited a similar pattern as the blot with anti-flagellin, except that the reduced STF2.HA1-2 protein had a diminished signal compared to that of the non-reduced sample, indicating that a substantial portion of the anti-HA polyclonal antibodies in the antiserum recognize conformational epitopes formed by correct disulfide-bond formation while a small proportion of the antibodies in the anti-sera recognize linear epitopes that are not reliant on disulfide bound formation. Similar were observed using a capture ELISA assay. It is worth noting that anti-serum against B/Malaysia/2506/2004 (Victoria lineage) also recognizes B/Shanghai/361/2002,

which belongs to Yamagata lineage. This result indicates that these two lineages not only share substantial primary sequence similarity, but also share tertiary structure.

The western blot and ELISA results confirm the expression of functional Influenza B HA1-1 and HA1-2 fusion proteins in *E. coli*.

EXAMPLE 4: PURIFICATION AND BIOCHEMICAL CHARACTERIZATION OF RECOMBINANT FLAGELLIN-HEMAGGLUTININ FUSION PROTEINS

Materials and Methods

Bacterial growth and cell lysis: HA constructs were expressed in the *E. coli* host strain BLR (DE3). *E. coli* cells carrying a plasmed encoding STF2.HA1-1.his(PR8) (SEQ ID NO: 54) were cultured and harvested as described above. Individual strains were retrieved from glycerol stocks and grown in shake flasks to a final volume of 12L. Cells were grown in LB medium containing 50 µg/ml kanamycin/ 12.5 µg/ml tetracycline/0.5% dextrose to OD_{600} = 0.6 and induced with 1 mM IPTG for 3 h at 37°C. The cells were harvested by centrifugation (7000 rpm x 7 minutes in a Sorvall RC5C centrifuge) and resuspended in 1X PBS, 1% glycerol, 1 µg/ml DNAseI, 1 mM PMSF, protease inhibitor cocktail and 1 mg/ml lysozyme. The cells were then lysed by two passes through a microfluidizer at 15,000 psi. The lysate was centrifuged at 45,000 g for one hour in a Beckman Optima L ultracentrifuge (Beckman Coulter; Fullerton, CA) to separate the soluble and insoluble fractions

Purification and refolding of STF2.HA1-1.his(PR8) (SEQ ID NO: 89) protein from *E. coli*. Following lysis and centrifugation, the insoluble (pellet) fraction was resuspended in 100 ml 1X TBS, pH 8.0 + 1 mM β-mercaptoethanol and solid urea was added to a final concentration of 6.2 M. After resuspension with a Dounce glass-ball homogenizer the mixture was centrifuged as described above. The resulting supernatant was loaded onto a chelating sepharose column (GE/Amersham Biosciences; Piscataway, NJ) charged with $NiSO_4$ and equilibrated in Buffer A [20 mM Tris, pH 8.0/8 M urea/0.5 M NaCl].

After washing with 1 L buffer B [Buffer A + 1% (w/v) TX-100] the column was eluted in a 5-column volume linear gradient from 100% Buffer A to 100% Buffer C [Buffer A + 0.5 M imidazole]. Peak fractions were pooled, concentrated

3.5 fold on an Amicon 15 spin concentrator (Millipore; Billerica, MA) and dialyzed against 3 x 2 L of 8 M urea/20 mM Tris, pH 8.0/2 mM EDTA. Following dialysis, STF2.HA1-1.his(PR8) (SEQ ID NO: 89) was refolded by rapid dilution to a final concentration of 0.1 mg/ml protein in Refolding Buffer [0.1 M Tris, pH 8.0/0.1 M NaCl/1%(w/v) glycerol/5 mM reduced glutathione/1 mM oxidized glutathione] and incubated overnight at room temperature.

The refolded protein was then captured on a 5 ml chelating sepharose HiTrap column (GE/Amersham Biosciences; Piscataway, NJ) charged with $NiSO_4$ and equilibrated in Buffer D [20 mM Tris, pH 8.0/0.5 M NaCl/1 mM reduced glutathione/0.2 mM oxidized glutathione] and eluted in 100% Buffer E [Buffer D + 0.5 M imidazole]. Peak eluate fractions were pooled, concentrated using an Amicon 15 spin concentrator (Millipore; Billerica, MA) and fractionated on a Superdex 200 size-exclusion column (GE/Amersham; Piscataway, NJ) equilibrated in 1X TBS, pH 8.0 + 0.5% (w/v) Na deoxycholate. S200 peak fractions were pooled, dialyzed against 3 x 2 L of 1X TBS, pH 8.0 (without deoxycholate), vialed and stored at -80°C.

Purification of STF2.HA1-2(PR8) (SEQ ID NO: 90) and STF2.HA1-2*mut*(PR8) (SEQ ID NO: 91): *E. coli* cells carrying a plasmid encoding STF2.HA1-2.his (PR8) (SEQ ID NO: 55) were cultured and harvested as described above. Following lysis and centrifugation, the insoluble (pellet) fraction containing STF2.HA1-2(PR8) was reuspended and homogenized in 50 mM Tris, pH 8.0 + 0.5% (w/v) Triton X-100 using a Dounce homogenizer. The homogenate was then centrifuged (19,000 rpm x 10 min). The resulting pellet fraction was then re-homogenized in 50 mM Tris, pH 8.0 + 0.5% (w/v) Triton X-100 + 1.0 M NaCl and centrifuged. The inclusion body fraction (pellet) was then dissolved in buffer A [50 mM Na Acetate, pH 4.0 + 8.0 M urea] and centrifuged (19,000 rpm x 10 min.) to remove insoluble debris.

The resulting supernatant was fractionated on a Source S column (GE/Amersham Biosciences; Piscataway, NJ) equilibrated in Buffer A and eluted in a 5 column-volume linear gradient with Buffer B (50 mM Na Acetate, pH 4.0, 1.0 M NaCl). The protein was then refolded by rapid dilution into Buffer C (100 mM Tris-HCl, pH 8.0). The refolded protein was then fractionated on a Q sepharose HP

column (GE/Amersham Biosciences; Piscataway, NJ) equilibrated in Buffer C and eluted in a 0 % to 60% linear gradient with buffer D (Buffer C + 1 M NaCl). The Q HP eluate was then fractionated on a Superdex 200 size exclusion chromatography (SEC) column (GE/Amersham Biosciences; Piscataway, NJ) equilibrated with 1X Tris-buffered saline, pH 8.0, to separate monomeric from aggregated protein. Monomeric fractions were then pooled, aliquoted and stored at -80°C.

Purification of STF2.HA1-3(PR8) (SEQ ID NO: 92): *E. coli* cells carrying a plasmed encoding STF2.HA1-3 (PR8) (SEQ ID NO: 57) were cultured and harvested as described above. Following lysis and centrifugation the soluble (supernatant) fraction containing STF2.HA1-3(PR8) was adjusted to 6.2 M urea, 100 mM DTT and incubated at room temperature overnight to completely reduce and denature the protein. The protein was then diluted five-fold in Buffer A (50 mM citric acid, pH 3.5, 8 M urea, 1 mM β-mercaptoethanol, 1 mM EDTA) and loaded onto an SP sepharose FF column (GE/Amersham Biosciences; Piscataway, NJ) equilibrated with Buffer A. To remove endotoxin, the column was washed with 10 column volumes of Buffer B (Buffer A + 1% (w/v) Triton X-100) followed by 10 column volumes Buffer C (50 mM citric acid, pH 3.5, 1 mM β-mercaptoethanol, 1 mM EDTA, 70%(v/v) isopropanol).

The bound protein was then eluted in a five column-volume linear gradient of Buffer A:Buffer D (Buffer A + 1M NaCl). The eluted protein was then dialyzed to Buffer E (20 mM Tris, pH 8.5, 8 M urea, 1 mM β-mercaptoethanol, 1 mM EDTA) and passed through a Source Q column (GE/Amersham Biosciences; Piscataway, NJ). The STF2.HA1-3(PR8) (SEQ ID NO: 92) protein was collected in the flow-thru fraction. The protein was concentrated using Amicon spin-concentrators (Millipore; Billerica, MA) and dialyzed to Buffer F (20mM Tris, pH 8.0, 8 M urea, 1 mM EDTA) to remove reductant. The denatured polypeptide was then refolded by rapid dilution to a final concentration of 0.1 mg/ml in Refolding Buffer (0.1 M Tris, pH 8.0, 0.1 M NaCl, 1%(w/v) glycerol, 5 mM reduced glutathione, 1 mM oxidized glutathione) and incubated overnight at room temperature. The refolded protein was then fractionated on a Superdex 200 size exclusion column equilibrated in 1X Tris-buffered saline (TBS), pH 8.0, plus 0.25% (w/v) sodium deoxycholate to separate monomeric STF2.HA1-3(PR8) from

aggregated protein. Monomeric peak fractions were pooled, dialyzed to 1X TBS, pH 8.0, to remove deoxycholate, aliquoted and stored at -80°C.

Purification of HA1-2His(PR8) (SEQ ID NO: 93): *E. coli* cells carrying a plasmed encoding STF2.HA1-2.his (PR8) (SEQ ID NO: 61) were cultured and harvested as described above. Following cell lysis the insoluble fraction was homogenized and washed once with Buffer A (50 mM Tris, pH 8.0), followed by sequential washing with Buffer A + 0.5% (w/v) Triton X-100 and Buffer A + 0.5% (w/v) Triton X-100 + 1M NaCl. The insoluble material was then washed twice more with Buffer A alone. The washed inclusion bodies were then dissolved in Buffer B (1X PBS, pH 7.4, 8 M urea) and applied to a Ni-NTA (Sigma-Aldrich; St. Louis, MO) column charged with $NiSO_4$ and eluted in a linear gradient of Buffer C (50 mM Tris, pH 8.0, 8 M urea, 0.3 M NaCl, 0.5 M imidazole). In order to reduce endotoxin, the eluted protein was dialyzed overnight against 1X PBS, pH 7.4, 8 M urea, and reapplied to the Ni-NTA column.

The column was then washed with Buffer D (10 mM Tris, pH 8.8, 8 M urea, 60% (v/v) isopropanol) and eluted as described above. Ni-NTA purified protein was then refolded by 1:10 (v/v) dilution into Refolding Buffer (50 mM Tris, pH 8.8). The refolded HA1-2 His(PR8) (SEQ ID NO: 93) protein was then concentrated using an Amicon ultrafiltration unit (Millipore; Billerica, MA) and applied to a Superdex 200 size exclusion column (GE/Amersham; Piscataway, NJ) equilibrated in 1X TBS, pH 8.0. Monomeric peak fractions were pooled, aliqoted and stored at -80°C.

SDS-PAGE and western blot analysis: Protein identity of all HA constructs was determined, and purity estimated, by SDS-PAGE. An aliquot of 5 µg of each sample was diluted in SDS-PAGE sample buffer with or without 100 mM DTT as a reductant. The samples were boiled for 5 minutes and loaded onto a 10% SDS polyacrylamide gel (LifeGels; French's Forrest, New South Wales, AUS) and electrophoresed. The gel was stained with Coomassie R-250 (Bio-Rad; Hercules, CA) to visualize protein bands. For western blot, 0.5 µg/lane total protein was electrophoresed as described above and the gels were then electro-transfered to a PVDF membrane and blocked with 5% (w/v) dry milk before probing with anti-flagellin antibody (Inotek; Beverly, MA) or influenza A PR/8/34 Convalescent

immune serum (described below under Protein Antigenicity ELISA). After probing with alkaline phosophatase-conjugated secondary antibodies (Pierce; Rockland, IL), protein bands were visualized with an alkaline phosphatase chromogenic substrate (Promega; Madison, WI).

Protein assay: Total protein concentration for all proteins was determined using the Micro BCA (bicinchoninic acid) Assay (Pierce; Rockford IL) in the microplate format, using bovine serum albumin as a standard, according to the manufacturer's instructions.

Endotoxin assay: Endotoxin levels for all proteins were determined using the QCL-1000 Quantitative Chromogenic LAL test kit (Cambrex; E. Rutherford, NJ) , following the manufacturer's instructions for the microplate method.

TLR5 bioactivity assay: HEK293 cells constitutively express TLR5, and secrete several soluble factors, including IL-8, in response to TLR5 signaling. Cells were seeded in 96-well microplates (50,000 cells/well), and the following test proteins were added and incubated overnight: STF2.HA1-1His(PR8) (SEQ ID NO: 89); STF2.HA1-2(PR8) (SEQ ID NO: 90); STF2.HA1-2(PR8)*mut* (SEQ ID NO: 91); STF2.HA1-3(PR8) SEQ ID NO: 92); HA1-2His(PR8) (SEQ ID NO: 93); STF2.HA1-2 (Mal) (SEQ ID NO: 211) supernatant; STF2. HA1-2 (Mal) (SEQ ID NO: 211) refolded; STF2.HA1-2 (SH) (SEQ ID NO: 211) supernatant; STF2.HA1-2 (SH) (SEQ ID NO: 211) refolded; STF2.HA1-1 (Mal) (SEQ ID NO: 209) supernatant; and STF2.HA1-1 (Mal) (SEQ ID NO: 209). The next day, the conditioned medium was harvested, transferred to a clean 96-well microplate, and frozen at -20°C. After thawing, the conditioned medium was assayed for the presence of IL-8 in a sandwich ELISA using an anti-human IL-8 matched antibody pair (Pierce, Rockland, IL; #M801E and #M802B) following the manufacturer's instructions. Optical density was measured using a microplate spectrophotometer (FARCyte, GE/Amersham; Piscataway, NJ).

Protein antigenicity ELISA: To determine whether the recombinant fusion proteins displayed correctly folded epitopes of HA, the antigenicity of individual HA-fusion proteins was evaluated by ELISA. 96-well ELISA plates were coated overnight at 4°C with serial dilutions in PBS (100 µl/well) of each target protein starting at 5 µg/ml. Plates were blocked with 200 µl/well of Assay Diluent Buffer

(ADB; BD Pharmingen) for one hour at room temperature, then washed three times in PBS-T. A fixed dose of primary antibody was then added to each well.

To assay HA reactivity, 100 µl/well of a 1:10,000 dilution of non-immune or PR/8/34 convalescent immune serum in ADB was added. PR/8/34 immune serum was generated in BALB/c mice (Jackson Laboratory, Bar Harbor, ME) that received an experimentally determined sublethal challenge dose of $8x10^1$ egg infectious dosages (EID) of PR/8/34 influenza virus. Animals were then allowed to convalesce for >21days post-infection at which time immune serum was isolated and clarified. For ELISA of flagellin or the 6x histidine tag, monoclonal antibody against 6x His (Invitrogen; Carlsbad, CA), or flagellin (Inotek; Beverly, MA) was added at 1 µg/ml in ADB (100 µl/well) and the plates were incubated for 1 hr at room temperature or overnight at 4°C. The plates were then washed three times with PBS-T. HRP-labeled goat anti-mouse IgG antibodies (Jackson Immunochemical; West Grove, PA) diluted in ADB were added (100 µl/well) and the plates were incubated at room temperature for 1 hour. The plates were washed three times with PBS-T. After adding TMB Ultra substrate (Pierce; Rockford, IL) and monitoring color development, A_{450} was measured on a microplate spectrophotometer (FARCyte, GE/Amersham; Piscataway, NJ).

Results and Discussion

Protein yield and purity: Results for the purification of recombinant HA and STF2.HA fusion proteins produced in *E. coli* are shown in Table 6. All four proteins were produced in high yield, with estimated purity exceeding 90% and endotoxin well below the standard acceptable level of 0.1 EU/µg. The three STF2 fusion proteins also had very high *in vitro* TLR5 bioactivity while the HA1-2His_6 (SEQ ID NO: 93) protein, as expected, had no TLR5 activity.

Table 6

protein	SEQ ID NO:	yield (mg)	purity est. (%)	endotoxin (EU/μg)	TLR5 activity (EC_{50}, ng/ml)
STF2.HA1-1His(PR8)	89	6.4	>90	0.04	0.2
STF2.HA1-2(PR8)	90	6.0	>98	0.02	0.15
STF2.HA1-2(PR8)*mut*	91	12	>95	0.01	1.2
STF2.HA1-3(PR8)	92	2.1	>90	0.016	1.2
HA1-2His(PR8)	93	10	>95	0.03	nd

Antigenicity of Influenza A HA proteins produced in *E. coli*.: All four recombinant proteins were analyzed by western blotting with antibody against STF2 (flagellin) (Inotek; Beverly, MA) and immune antisera collected from PR/8/34 convalescent mice. The three STF2 fusion proteins appeared to react comparably with anti-flagellin antibody. STF2.HA1-1His(PR8) (SEQ ID NO: 89), STF2.HA1-2(PR8) (SEQ ID NO: 90) and HA1-2His(PR8) (SEQ ID NO: 93) also reacted with anti-PR/8/34 convalescent serum. This reactivity was only seen in non-reduced protein samples; the addition of DTT to the sample buffer greatly diminished the signal. This finding suggests that the majority of the anti-HA antibodies in the convalescent serum are conformationally dependent and that proper disulfide bonding in HA is necessary for reactivity by western blot. In contrast to these results, STF2.HA1-3(PR8) (SEQ ID NO: 92) did not react efficiently with the convalescent serum and what little signal was produced was not affected by reductant. This finding suggested that the STF2.HA1-3 protein may not be capable of folding properly and/or presenting native HA epitopes.

Capture ELISA for Influenza B Globular Head Proteins

The antigenicity of individual HA-fusion proteins was evaluated using capture ELISA. An anti-flagellin monoclonal antibody was used on the solid phase to capture these proteins. Ferret anti-sera raised on natural B/Malaysia/2506/2004 infection was used to detect the captured HA-fusion proteins. The results of this assay will determine whether the recombinant fusion proteins displayed correctly-folded epitopes of HA. Due to the considerable sequence homology between the

different Influenza B HA globular head molecules, 84.6%, these antisera would be expected to react with both the Malaysian and the Shanghai globular head constructs.

Preparation of Samples

E. coli cells expressing the STF2.HA1-2(Malaysia (SEQ ID NO: 210)), STF2.HA1-2(Shanghai (SEQ ID NO: 211)), and STF2.HA1-1 (Malaysia (SEQ ID NO: 209)) recombinant proteins were pelleted by centrifugation and resuspended in Tris buffer (50mM Tris, 200mM NaCl, pH 8). Cells were lysed by sonication. The cell lysate was centrifuged to separate the soluble proteins from the insoluble proteins. The insoluble protein pellet was resuspended in Tris buffer with 6M urea and the proteins were subjected to rapid refolding by quickly diluting the solution 1:10 fold in Tris buffer. The protein concentrations of both the soluble proteins and the pelleted samples containing refolded proteins were estimated by UV_{280} (Spectrophotometer DU 800). Supernatants containing STF2.HA1-1(Mal) (SEQ ID NO: 209); STF2.HA1-2(Mal) (SEQ ID NO: 210), STF2.HA1-2(SH, Shanghai) (SEQ ID NO: 211) protein and solubilized pellets containing refolded STF2.HA1-1(Mal), STF2.HA1-2(Mal), STF2.HA1-2(SH) were evaluated by ELISA for reactivity with ferret antisera raised on natural infection against B/Malaysia/2506/2004. *E. coli* supernatant samples and solubilized, refolded samples containing STF2.4xM2e (SEQ ID NO: 94)protein were used as negative controls.

ELISA Method

The ELISA plate (Maxisorp, Nunc, Denmark) was coated with the flagellin-specific monoclonal antibody 6H11 (Inotek, Beverly, MA) at 0.5ug/mL and incubated overnight at 2-5°C. The antibody coating solution was aspirated and the wells were blocked with 300uL/well Super Block + Tween-20 for 2 hours at 25°C. The plate was washed once with 1xPBS and blot-dried. Three-fold serial dilutions of the different protein solutions were performed in a dilution plate starting with a concentration of 5ug/mL. 100uL was transferred from the dilution plate to the ELISA plate.

The plate was incubated for 1hr at 25°C. Unbound protein was removed by washing the plate 3 times with PBS+0.05%Tween-20. Ferret antiserum raised against B/Malaysia/2506/2004 (CDC, AL, Georgia) was diluted to 1:100 and 100uL was added to each well. The plate was then incubated for 1hr at 25°C. After this incubation step, the plates were washed 6 times with PBS containing 0.05% Tween-20. Goat anti-ferret IgG conjugated to horseradish peroxidase (HRP, Bethyl Labs Inc., IL) was diluted 1:10,000 and 100uL was added to each well. The plate was incubated for 30 min. After this incubation step, the plate was washed 6 times with PBS+0.05%Tween-20. 100uL of TMB Ultra containing the HRP substrate 3,3',5,5'-tetramethylbenzidine (Pierce, Rockford, IL) was added to each well. After the addition of this substrate, the color development was monitored and the reaction was stopped with the addition of 100uL/well 1M H_2SO_4 . A_{450} was measured on a microplate reader (SpectraMax 190, Molecular devices, Sunnyvale, CA).

Antigenicity of Influenza B HA Proteins Produced in *E. coli*

The ELISA data (Figure 3) indicated that the antisera recognized epitopes that were sensitive to reductant. Treatment of proteins with reducing agents, which alter the conformation of the proteins by the disruption of disulphide bonds and, diminished reactivity of the antisera with the proteins. The reactivity of STF2.HA1-2(Mal) total lysate non-reduced, STF2.HA1-2(SH) total lysate non-reduced and STF2.HA1-1(Mal) total lysate non-reduced when maintained in the properly folded configuration was very good and indicates that the conformation of the individual HA-fusion proteins was comparable to that of the globular head in the native HA protein. While the antiserum contains conformationally sensitive antibodies and the reactivity depends on correct disulfide bonding, it also contains antibodies that are non-sensitive to disulfide bonding as demonstrated by the residual activity which was observed in both reduced and non-reduced conditions.

TLR5 Bioactivity of STF2.HA (Influenza B) Proteins

The recombinant fusion proteins showed potency which matched the results from the Capture ELISA indicating that refolded samples may be more active than the lysate in its native form (Figures 4 and 5). The samples in the native untreated form showed activity which were 2-fold less than the refolded samples. These

proteins were misfolded in the native form and therefore require a refolding step to reinstall the TLR-5 bioactivity. This activity was consistent with HA activity in the Capture ELISA.

EXAMPLE 5: IMMUNOGENICITY OF RECOMBINANT FLAGELLIN-HEMAGGLUTININ FUSION PROTEINS REPRESENTING VIRAL STRAIN A/PUERTO RICO/8/34 in *E. coli*

Materials and Methods

Animal studies: Female BALB/c mice (Jackson Laboratory, Bar Harbor, ME)) were used at the age of 6-8 weeks. Mice were divided into groups of 10 and received inguinal subcutaneous (s.c) immunizations on days 0 and 14 as follows:

1) PBS (phosphate buffered saline).
2) 3 μg of STF2.4xM2e (SEQ ID NO: 94) in saline buffer (10mM Histidine, 10mM Tris, 75mM NaCl, 5% (vol/vol) sucrose, 0.02% (w/v) Polysorbate-80, 0.1mM EDTA, 0.5% (v/v) ethanol, pH 7.2)
3) 30 μg of STF2.HA1-2(PR8) (SEQ ID NO: 90) in PBS
4) 3.0 μg of STF2.HA1-2(PR8) (SEQ ID NO: 90) in PBS
5) 0.3 μg of STF2.HA1-2(PR8) (SEQ ID NO: 90) in PBS
6) 0.03 μg of STF2.HA1-2(PR8) (SEQ ID NO: 90) in PBS

An additional group of five mice received an experimentally determined sublethal challenge with $8x10^1$ egg infectious dosages (EID) PR/8/34 and were allowed to convalesce for >21days. These animals were then used as immune convalescent positive controls during the challenge studies (see below). Mice were bled on days 10 (primary) and 21 (boost), and sera were clarified by clotting and centrifugation and stored at -20°C.

Serum antibody determination: HA-specific IgG levels were determined by ELISA. 96-well ELISA plates (Costar (Cat #9018) Corning, NY) were coated overnight at 4°C with 100 μl/well HA0sHis protein (produced in *Drosophila*) (SEQ ID NO: 176) in PBS (5 μg/ml). Plates were blocked with 200 μl/well of Assay Diluent Buffer (ADB; BD Pharmingen, (Cat#: 555213)(San Diego, CA) for one hour at room temperature. The plates were washed three times in PBS+ 0.05% (v/v) Tween 20 (PBS-T). Dilutions of the sera in ADB were added (100 μl/well) and the

plates were incubated overnight at 4°C. The plates were washed three times with PBS-T. HRP-labeled goat anti-mouse IgG antibodies (Jackson Immunochemical, West Grove, PA (Cat#: 115-035-146)) diluted in ADB were added (100 µl/well) and the plates were incubated at room temperature for 1 hour. The plates were washed three times with PBS-T. After adding TMB Ultra substrate (Pierce (Cat 34028), Rockford, IL)) and monitoring color development, A_{450} was measured on a Tecan Farcyte (Durham, NC) microplate spectrophotometer.

MDCK whole cell ELISA: MDCK cells (ATCC (Cat# CCL-34) Manassas, VA) were grown in 96-well culture plates (BD (Cat 353075), Corning, NY) in DMEM complete medium containing 10% FCS at 37°C for one to two days or until cells were near confluence. Wells were then incubated with $1x10^6$ EID of PR/8/34 virus (50 µl) in DMEM medium without FCS or medium alone (for uninfected controls).

Following a 60 minute incubation at 37°C, 200 µl of complete medium was added to each well and plates were incubated overnight at 37°C. The next day plates were washed with PBS and fixed with 10% formalin at room temperature for 10 minutes. Wells were washed three times with PBS/0.1% BSA and blocked with 200 µl/well ADB (BD Pharmingen, Cat# 555213), San Diego, CA)for one hour at RT or overnight at 4°C. Serial dilutions of test sera were added to the wells and incubated for one to two hours at room temperature. Wells were washed and incubated with HRP-conjugated goat anti-mouse IgG (Jackson Immunochemical, Cat 115-035-146 (West Grove, PA) for 30 minutes at room temperature, followed by TMB Ultra substrate (Pierce (Cat# 34028), Rockford, IL) for two minutes at room temperature. The reaction was stopped with the addition of 25 µl of 1N H_2SO_4 and the OD_{450} was read using a microplate spectrophotometer (FARCyte, Amersham, Durham, NC).

Results and Discussion

Induction of HA-specific IgG response following immunization with STF2.HA1-2(PR8) (SEQ ID NO: 90): The immunogenicity of STF2.HA1-2 was examined by immunizing BALB/c mice (10/group) subcutaneously on day 0 and 14 with a dose range of STF2.HA1-2 (30, 3.0, 0.3, or 0.03 µg). Control groups of mice were immunized with PBS (negative control), 3 µg of STF2.4xM2e (SEQ ID NO:

94) (negative control for HA immunogenicity, positive control for lethal challenge efficacy study), or a sub-lethal challenge with $8x10^1$ EID of the influenza isolate PR/8/34 to generate immune convalescent animals. HA-specific IgG responses were examined 7 days post boost (Day 21) by ELISA. The results demonstrate that immunization with 30, 3 or 0.3 μg of STF2.HA1-2(PR8) induced consistent and significant HA0sHis-specific (SEQ ID NO: 176) IgG responses in a dose-dependent manner (Figure 6).

Sera from BALB/c mice immunized with STF2.HA1-2 (SEQ ID NO: 90) react with MDCK cells infected with influenza virus: The direct ELISA results above demonstrated that the immune sera from STF2.HA1-2-immunized animals recognize recombinant *Drosophila* expressed HA0sHis (SEQ ID NO: 176) corresponding to the PR/8/34 HA sequence. In order to demonstrate that the anti-HA antibodies recognize native viral HA, the same sera were examined for the ability to react with MDCK cells infected with PR/8/34. The results shown in Figure 7 demonstrate that sera from mice immunized with 30, 3.0, or 0.30 μg of STF2.HA1-2 bound specifically to PR/8/34-infected MDCK cells, indicating that the anti-HA antibodies elicited by immunization with STF2.HA1-2 recognize HA in its native conformation.

Collectively these results demonstrate that mice immunized with STF2.HA1-2(PR8) (SEQ ID NO: 90) in PBS, without conventional adjuvant or carrier, mounted vigorous anti-HA responses that recognized recombinant HA expressed in *Drosophila* and native viral HA expressed on the surface of PR/8/34 infected MDCK cells *in vitro*.

EXAMPLE 6: EFFICACY OF RECOMBINANT FLAGELLIN-HEMAGGLUTININ FUSION PROTEINS REPRESENTING VIRAL STRAIN A/PUERTO RICO/8/34

Materials and Methods

Influenza virus challenge of mice. To assess efficacy, mice immunized as described in Example 5 were challenged on day 28 by intranasal administration of an LD_{90} (dose lethal to 90% of mice)($8x10^3$ EID) of influenza A isolate PR/8/34. Animals were monitored daily for 21 days following the challenge for survival, weight loss and clinical presentation. The % weight loss was calculated based on

the mean of ((Daily weight (g) /Initial weight (g) day 28) x 100) of each individual animal per group. Clinical scores were assigned as follows: 4 pts=healthy, 3 pts= reduced grooming, 2 pts= reduced physical activity and 1 pt = moribund. (Experimental results for clinical scores and weight loss reflect the results based on surviving animals on the day evaluated).

Virus titration and determination of $TCID_{50}$.

Cell preparation: MDCK cells (ATCC (Cat# CCL-34), Manassas, VA) were cultured in 100x20mm culture plates (BD (Cat# 353003) Corning, NY) to 90-95% confluency and the monolayer was dislodged by incubation with trypsin-EDTA at 37°C/5% CO_2 for 20 minutes. The trypsin was inactivated by the addition of DMEM cell culture medium supplemented with 10% FBS, and the cell monolayer was scraped with a sterile spatula to complete detachment of cells. The cell suspension was harvested and washed twice with DMEM+10%FBS. Cells were resuspended in DMEM+10% FBS and counted. The cell concentration was adjusted to 4 x 10^5 cells/ml and 100 µl was added to each well of a 96-well tissue culture plate (BD (Cat# 353075), Corning NY). Plates were incubated at 37°C/5%CO_2 until confluence reached 90-95%.

Viral titration: Influenza virus (strain A/Puerto Rico/8/34 [PR/8]) was diluted to 1x10^8 EID in phenol red-free DMEM+0.1%BSA (fractionV (Rockland (Cat#BSA-50)Gilbertsville, PA), and serial 5-fold dilutions were prepared in 96-well plates using the same medium. Monolayers of MDCK cells in 96-well plates, prepared as described in herein, were washed by aspirating the culture medium, replacing with 200 µl/well of 1x PBS, and aspirating the PBS. Serial 5-fold dilutions of influenza virus prepared above were added to the washed monolayers in a volume of 100 µl/well. One row of wells was treated with medium only as a control. The cells were incubated for 2 hours at 37°C/5% CO_2 to allow viral attachment and entry. Wells were washed by aspiration, rinsing with 200 µl/well PBS, and aspiration of the PBS. All wells received 100 µl/well phenol red-free DMEM+0.1% BSA and the plates are incubated for 48 hours at 37°C/5% CO_2.

Determination of cell viability: Following incubation with virus as described above, the medium was aspirated from the wells and replaced with fresh medium containing 40 µg/ml neutral red (Simga Aldrich (Cat# N2889) St. Louis, MO). To

determine maximum lysis, 2 µl lysis solution (9% Triton X-100 in water, weight/vol) was added to triplicate wells that were incubated with medium only. Following a one hour incubation, the cells were fixed by the addition of 100 µl/well 1% formaldehyde/1% $CaCl_2$ for 5 minutes at room temperature; this fix step was performed twice in succession. The fix solution was aspirated and the neutral red was released by the addition of 100 µl/well of extraction medium (50% ethanol/1% acetic acid).

The plate was incubated at room temperature for 20 minutes, with shaking for the final 2 minutes. The amount of dye released was determined by measuring absorbance at a wavelength of 540 nm using a microplate spectrophotometer. Cell death (and hence, viral infectivity) was measured as a decrease in the amount of dye released as compared to media control. At the end of the incubation, the neutral red assay was performed as described herein. The percentage lysis of each serum dilution was calculated as:

$$\% \text{ reduction} = 100 \times ((\text{sample-virus})/(\text{med-virus}))$$

where sample, max, and med refer to the absorbance values in wells representing experimental samples, virus only, and medium only, respectively. The neutralizing titer of each sample was defined as the dilution of serum which resulted in a 50% reduction in viral infectivity.

Influenza A/PR/8/34 Neutralization Assay This assay was adapted with modifications from WHO Manual on Animal Influenza Diagnosis and Surveillance, p. 86-88 (WHO/CDS/CSR/NCS2002.5). The Neutral Red Assay is adapted from a public protocol of the Cell Lab at Gettysburg College (http://www3.gettysburg.edu/~sorense/Celllab04/neutralred.htm)

Specifically, MDCK (ATCC, Cat#(CCL_34), Manassas, VA) cells were plated in 96-well tissue culture plates (BD (Cat 353075), Corning, NY) as described above. Test reagents (experimental and control sera) were heat-inactivated by incubating for 30 minutes in a water bath heated to 56°C. Sera were serially titrated in 96-well plates, in 3-fold dilutions in phenol red-free DMEM+0.1% BSA. An equal volume of PR/8/34 virus diluted to $5x10^6$/ml EID in the same medium was added to each serum dilution to achieve a final viral concentration of $2.5x10^6$

EID/ml (The pre-determined $TCID_{50}$ for our current stock of virus). Wells containing medium only and virus only were included as negative and positive controls, respectively. The plates were incubated for 30 minutes at 37°C/5% CO_2.

Cell monolayers prepared as described above (Cell preparation) were washed once with 200 µl/well 1x PBS, then 100 µl/well of serum:virus mixtures and control reagents prepared as described above were added and incubated for 2 hours at 37°C/5% CO_2. Following incubation, cell monolayers were washed once with PBS, and 100 µl/well phenol red-free DMEM+0.1% BSA was added to each well and incubated at 37°C/5% CO_2 for 48 hours. At the end of the incubation, the neutral red assay was performed as described above (Determination of cell viability). The percentage lysis of each serum dilution was calculated as described above. The neutralizing titer of each sample was defined as the dilution of serum which results in a 50% reduction in viral infectivity

Results and Discussion

Immunization with STF2.HA1-2(PR8) (SEQ ID NO: 90) provides protection from a lethal challenge with influenza A: The results from Example 5 demonstrated that immunization of mice with STF2.HA1-2(PR8) (SEQ ID NO: 90) generated an antibody response that recognized native HA. In order to evaluate efficacy, the same mice were challenged on day 28 with an LD_{90} ($8x10^3$EID) of PR/8/34 virus administered intra-nasally. Mice were monitored daily for 19 days following the challenge for survival, weight loss and clinical presentation. As shown in Figures 8A, 8B, and 8C, PBS-immunized mice showed signs of infection (weight loss and lower clinical scores) as early as three days post-challenge and all mice died by day 21 post-challenge. In contrast, mice immunized with 30, 3.0 or 0.3 µg of STF2.HA1-2(PR8) (SEQ ID NO: 90) demonstrated markedly enhanced protection. These animals demonstrated little to no weight loss, significantly higher clinical scores and 100% survival that was similar to immune convalescent control animals over the 21 day period. These results demonstrate that *E. coli*-expressed STF2.HA1-2(PR8) (SEQ ID NO: 90) induces HA-specific immune responses that successfully protect BALB/c mice from a lethal challenge with virulent influenza A virus *in vivo*.

In order to evaluate the effectiveness of the antibody response in vitro a cell-based virus neutralization assay was developed to test the ability of immune sera to neutralize viral infectivity. In this study the viral inhibitory activity of serum from animals immunized with STF2.HA1-2(PR8) (SEQ ID NO: 90) and STF2.HA1-1(PR8) (SEQ ID NO: 151) was examined. Serial dilutions of non-immune and immune sera were pre-incubated with PR/8 and incubated with MDCK cells. Wells were washed to remove free virus and plates were incubated prior to staining with neutral red to detect live cells. The results indicate that immune sera from animals immunized with STF2.HA1-2(PR8) (SEQ ID NO: 90) demonstrated a tissue culture inhibition 50% dose (TCID) of >1:40 (Figure 9).

EXAMPLE 7: IMMUNOGENICITY OF RECOMBINANT FLAGELLIN-HEMAGGLUTININ FUSION PROTEINS REPRESENTING VIRAL STRAIN A/VIET NAM/1203/2004

Materials and Methods

Animal studies: Female BALB/c mice (Jackson Laboratory, Bar Harbor, ME) were used at the age of 6-8 weeks. Mice were divided into groups of 10 and received inguinal s.c immunizations on days 0 and 14 as follows:

1) PBS (phosphate buffered saline).
2) 3.0 μg of STF2.HA1-2(PR8) (SEQ ID NO: 90) in PBS
3) 3.0 μg of STF2.HA1-2(VN) (SEQ ID NO: 95) in PBS
4) 0.3 μg of STF2.HA1-2(VN) (SEQ ID NO: 95) in PBS

Mice were bled on day 21 and sera were clarified by clotting and centrifugation and stored at -20°C.

Serum antibody determination: HA and STF2-specific IgG levels were determined by ELISA. 96-well ELISA plates (Costar (Cat#9018), Corning, NY) were coated overnight at 4°C with 100 μl/well of HA protein purified from Viet Nam/1203/2004 (BEI Resources (Cat #NR-660), Manassas, VA)) or recombinant flagellin (STF2) (SEQ ID NO: 96) in PBS (5 μg/ml). Plates were blocked with 200 μl/well of Assay Diluent Buffer (ADB; BD Pharmingen (Cat#555213) San Diego, CA) for one hour at room temperature. The plates were washed three times in PBS buffer containing 0.05% (v/v) Tween 20 (PBS-T). Dilutions of the sera in ADB

were added (100 µl/well) and the plates were incubated overnight at 4°C. The plates were washed three times with PBS-T. HRP-labeled goat anti-mouse IgG antibodies (Jackson Immunochemical (Cat#115-035-146), West Grove, PA) diluted in ADB were added (100 µl/well) and the plates were incubated at room temperature for 1 hour. The plates were washed three times with PBS-T. After adding TMB Ultra substrate (Pierce (Cat#34028), Rockford, IL)) and monitoring color development, A_{450} was measured on a Tecan Farcyte (Durham, NC) microplate spectrophotometer.

Results and Discussion

The immunogenicity of STF2.HA1-2(VN) (SEQ ID NO: 95) was examined in BALB/c mice. Animals were immunized s.c. on days 0 and 14 with 3 µg STF2.HA1-2(PR8) (SEQ ID NO: 90), or 3 or 0.3 µg of STF2.HA1-2(VN) (SEQ ID NO: 95). On day 21 animals were bled and serum IgG responses to purified HA from A/Viet Nam/1203/2004 (obtained from BEI Resousrces Cat# NR-660) and recombinant flagellin (STF2; (SEQ ID NO: 96)) were examined by ELISA (Figures 10A and 10B). The results demonstrate that serum from mice immunized with STF2.HA1-2(VN) (SEQ ID NO: 95) exhibit antigen-specific reactivity with purified H5 protein from influenza A/Viet Nam/1203/2004.

EXAMPLE 8: CLONING, EXPRESSION, AND BIOCHEMICAL CHARACTERIZATION OF RECOMBINANT FLAGELLIN-HEMAGGLUTININ FUSION PROTEINS PRODUCED IN *BACULOVIRUS*

Cloning: Expression of HA and STF2.HA fusion proteins was carried out using the Bac-to-Bac *Baculovirus* Expression System (Invitrogen, Carlsbad, CA) according to the manufacturer's guidelines (Figure 11). cDNAs encoding HA and STF2.HA fusion proteins were cloned into the pFastBac donor plasmid which was then used to transform DH10Bac cells containing the bacmid and helper DNA. Recombinant bacmid clones generated by homologous recombination in DH10Bac were then identified by blue-white screening on X-gal plates. The details of the cloning procedure are given in the following paragraphs.

Generation of vector cassettes: The pFastBac™1 vector is compatible with the Bac-to-Bac® *Baculovirus* Expression System (Invitrogen, Carlsbad, CA). This

vector has a strong AcMNPV polyhedrin (PH) promoter for high level protein expression and a large multiple cloning site for simplified cloning. pFastBac™1 is a non-fusion vector (*i.e.* no fusion tags are present in the vector).

To ensure proper expression of recombinant protein, the insert must contain an ATG start codon for initiation of translation and a stop codon for termination of translation. In order to facilitate cloning of flagellin in fusion with several truncations of the HA gene, two plasmid cassettes were engineered: pFB-STF2Blp.wt (SEQ ID NO: 97) and pFB-STF2Blp.ng (SEQ ID NO: 98). In the latter, the putative glycosylation sites in the flagellin (STF2) gene were obliterated by substituting Gln (Q) for Asn (N). Potential glycosylation sites were determined using the consensus sequence for N-glycosylation N-X-S/T: N residue followed by any residue by S or T in the third position.

In the certain locations, residues 40, 122, 215, 237, 414, 478, and 497 of the STF2 gene (SEQ ID NO: 212) the N residue was incorporated as a Q substitution. Constructs which harbor these mutations and in which the proteins are not likely to be modified by sugars are designated as non-glycosylated (ng) mutants. Similarly, glycosylation mutations were introduced in HA0s gene at positions 11, 23, 268, 286 and 480 (SEQ ID NO: 23). Both vector constructs (pFB-STF2Blp.wt and pFB-STF2Blp.ng) harbor a silent mutation at nucleotides 5'-GTGCTGAGCCTGTTACGT-3' (SEQ ID NO: 310) (nt 1501 to 1518) of STF2 (SEQ ID NO: 212) creating a unique *Blp*I in the the plasmid cassette. Each cassette contains the honey bee melittin (HBM) sequence (SEQ ID NO: 99) fused at the amino terminus of STF2 to provide a signal for secretion. Both plasmids were codon-optimized for expression in B*Baculovirus* (Midland Certified Reagent Co., Inc, Midland, TX; DNA2.0 Inc, Menlo Park, CA). The synthetic genes were excised with *Bam*HI or *Bgl*II and *Sph*I and cloned into pFastBac™1 vector by compatible end ligation generating the pFB-STF2Blp.wt and pFB-STF2Blp.ng cassettes.

Flagellin-HA fusion constructs: Subunits of the HA globular head from influenza strains A/Puerto Rico/8/34, A/Viet Nam/1203/2004, A/Indonesia/2005 and A/New Caledonia/12/99 were expressed alone or genetically fused to flagellin (STF2) (SEQ ID NO: 178) and expressed in *Baculovirus*. This was accomplished by one of three methods.

Method #1: To generate reagents for ELISA, HA1-1 subunits of PR8, VN, IND and NC strains (Tables 7 and 8) were cloned into pFASTBac™1, generating pFB.HA1-1 that harbors a hexa-his tag. This was achieved by employing PCR with a set of primers as indicated in the Tables below and with a synthetic codon-optimized HA0s gene (HA excluding the signal sequence and extracellular domain) as DNA template (DNA2.0 Inc., Menlo Park, CA). The PCR fragments were digested with *BglII* and *Sph*I and inserted into pFastBac™1 that has previously been treated with *BglII* and *Sph*I enzymes followed by BAP treatment. In order to generate a flagellin-HA subunit fusion, the PCR product was digested with *Blp*I, and ligated by compatible ends to pFB-STF2Blp.wt.

Method #2: In this protocol, codon-optimized HA subunit genes representing both the wt and non-glycosylated forms were chemically synthesized (DNA2.0 Inc., Menlo Park, CA). The genes were excised with *Blp*I and *Sph*I enzymes and the fragment gel purified and ligated to pFB-STF2Blp.wt or the non-glycosylated version, pFB-STF2Blp.ng previously digested with *Blp*I and *Sph*I and BAP treated. In each case, the ligation mix was used to transform TOP10 cells and transformants were screened by PCR and DNA sequencing to confirm the presence and correct orientation of the inserts. The constructs were used to transform MAX Efficiency® DH10Bac™ competent *E. coli* (Invitrogen, Carlsbad, CA) to generate a recombinant bacmid. The colonies of bacteria were screened for positive bacmids by blue/white selection on plates containing 50 μg/mL kanamycin, 7 μg/mL gentamycin, 10 μg/mL tetracycline, 40 μg/mL IPTG and Bluo-Gal (40 μg/mL). Recombinant bacmid DNA was prepared and used to transfect the insect cell line of choice (Sf9 or Sf21 cells) to generate a recombinant *Baculovirus*. The baculoviral stock was then amplified and titered and used to infect High Five insect cells to express the recombinant protein.

Table 7: PR8 HA constructs for expression in *Baculovirus*

SEQ ID NO:	Construct (His tag)	Method	FOR Primer SEQ ID NO:	REV Primer SEQ ID NO:	DNA Template SEQ ID NO:
100	STF2.HA0s	#1	101	102	103
104	STF2.HA1-1	#1	105	106	103
107	STF2.HA1-2	#1	108	109	103
110	STF2.HA1-2*mut*	#2	N/A	N/A	111
112	STF2.HA1-3	#1	113	114	103
115	STF2.HA1-3*mut*	#2	N/A	N/A	116
117	ngSTF2.HA0s	#2	N/A	N/A	118
119	ngSTF2.HA1-1	#2	N/A	N/A	120
121	ngSTF2.HA1-2	#2	N/A	N/A	122
123	ngSTF2.HA1-2*mut*	#2	N/A	N/A	124
125	ngSTF2.HA1-3	#2	N/A	N/A	126
127	ngSTF2.HA1-3*mut*	#2	N/A	N/A	128
129	wtSTF2.HA1-1ng	#2	N/A	N/A	130
131	ngSTF2.HA1-1wt	#2	N/A	N/A	132
133	HA1-1	#1	134	135	136
137	HA1-1 (no tag)	#1	134	138	136

Table 8: NC, VN and NC HA constructs for expression in *Baculovirus*

SEQ ID NO:	Construct (his tag)	Method	FOR Primer SEQ ID NO:	REV Primer SEQ ID NO:	DNA Temlat SEQ ID NO:
139	HA1-1(NC)	#1	140	141	142
143	HA1-1(VN)	#1	144	145	146
147	HA1-1(IND)	#1	148	149	150

Protein Expression: *Spodoptera frugiperda* (Sf9) insect cells were cultured in Grace's Insect Cell Medium (Invitrogen; Carlsbad, CA). Recombinant bacmids were purified from DH10Bac clones and transfected into Sf9 cells using Cellfectin (Invitrogen; Carlsbad, CA) to generate passage one (P1) *Baculovirus* stocks which

were then titered by traditional plaque assay according to the manufacturer's guidlines. Subsequent infections of Sf9 cells were performed using P1 viral stocks at a multiplicity of infection (MOI) of 0.1 to produce larger volume P2 viral stocks with increased viral titers. Cultures expressing the following proteins were scaled up for protein purification: ngSTF2.HA(PR8)1-2*mut*His (SEQ ID NO: 156); ngSTF2.HA(PR8)1-1His (SEQ ID NO: 152 ; STF2.HA(PR8)1-1His (SEQ ID NO: 151); STF2.HA(PR8)1-2His (SEQ ID NO: 153); STF2.HA1-2mutHis (SEQ ID NO: 155); HA(PR8)1-1His (SEQ ID NO: 179) HA(VN)1-1His (SEQ ID NO: 180); HA(IND)1-1His (SEQ ID NO: 181); HA(NC)1-1His (SEQ ID NO: 182). Protein expression was performed by infecting 2 x 1 L flasks of High-5 insect cells with P2 virus at an MOI of 2 and culturing for 24 hrs at 28°C with slow shaking. Conditioned medium was harvested by centrifugation and clarified by passing through a 0.22 μm filter and stored at 4°C.

Protein Purification: $NiSO_4$ was added to each clarified supernatant to a final concentration of 0.5 mM and the pH adjusted to 8.0. The His_6-tagged protein was then captured on a chelating sepharose column (GE/Amersham Biosciences; Piscataway, NJ) charged with $NiSO_4$ and equilibrated in Buffer A (20 mM Tris, pH 8.0, 0.5 M NaCl). After washing with Buffer A the bound protein was eluted with Buffer B (Buffer A + 0.5 M imidazole). Peak fractions were pooled, dialyzed overnight into Buffer A and applied to a nickel NTA column (Sigma-Aldrich; St. Louis, MO). The column was eluted with a 5-column volume gradient from 100% Buffer A to 100% Buffer B. Peak fractions were pooled and dialyzed into 1X TBS, pH 8.0.

SDS-PAGE and western blot analysis: Protein identity was determined, and purity estimated, by SDS-PAGE. An aliquot of 5 μg of each sample was diluted in SDS-PAGE sample buffer with or without 100 mM DTT as a reductant. The samples were boiled for 5 minutes and loaded onto a 10% SDS polyacrylamide gel (LifeGels; French's Forrest, New South Wales, AUS) and electrophoresed. The gel was stained with Coomassie R-250 (Bio-Rad; Hercules, CA) to visualize protein bands. For western blot, 0.5 μg/lane total protein was electrophoresed as described above and the gels were then electro-transfered to a PVDF membrane and blocked with 5% (w/v) dry milk before probing with anti-flagellin antibody (Inotek; Beverly,

MA), anti-His_6 antibody (Invitrogen; Carlsbad, CA), or influenza A PR/8/34 convalescent immune serum (described below under Protein Antigenicity ELISA). After probing with alkaline phosophatase-conjugated secondary antibodies (Pierce; Rockford, IL) , protein bands were visualized with an alkaline phosphatase chromogenic substrate (Promega; Madison, WI).

Protein assay: Total protein concentration was determined using the Micro BCA (bicinchoninic acid) Assay (Pierce; Rockford, IL) in the microplate format, using bovine serum albumin as a standard, according to the manufacturer's instructions.

Endotoxin assay: Endotoxin levels were determined using the QCL-1000 Quantitative Chromogenic LAL test kit (Cambrex; E. Rutherford, NJ) , following the manufacturer's instructions for the microplate method.

TLR5 bioactivity assay: HEK293 cells constitutively express TLR5, and secrete several soluble factors, including IL-8, in response to TLR5 signaling. Cells were seeded in 96-well microplates (50,000 cells/well), and the following test proteins were added and incubated overnight: ngSTF2.HA(PR8)1-2*mut*His (SEQ ID NO: 156); ngSTF2.HA(PR8)1-1His (SEQ ID NO: 152) ; STF2.HA(PR8)1-1His (SEQ ID NO: 151); STF2.HA(PR8)1-2His (SEQ ID NO: 153); STF2.HA1-2(PR8)*mut*His (SEQ ID NO: 155) The next day, the conditioned medium was harvested, transferred to a clean 96-well microplate, and frozen at -20°C. After thawing, the conditioned medium was assayed for the presence of IL-8 in a sandwich ELISA using an anti-human IL-8 matched antibody pair (Pierce; Rockford, IL, #M801E and #M802B) following the manufacturer's instructions.

Protein antigenicity ELISA: The following constructs purified from *Baculovirus* culture were evaluated by ELISA to determine whether the proteins displayed correctly folded epitopes of HA: ngSTF2.HA(PR8)1-2*mut*His (SEQ ID NO: 156); ngSTF2.HA(PR8)1-1His (SEQ ID NO: 152) ; STF2.HA(PR8)1-1His (SEQ ID NO: 151); STF2.HA(PR8)1-2His (SEQ ID NO: 153); STF2.HA1-2*mut*His (SEQ ID NO: 155); and HA(PR8)1-1His (SEQ ID NO: 179). 96-well ELISA plates were coated overnight at 4°C with serial dilutions in PBS (100 µl/well) of each target protein starting at 5 µg/ml. Plates were blocked with 200 µl/well of Assay Diluent Buffer (ADB; BD Pharmingen) for one hour at room temperature, then

washed three times in PBS-T. A fixed dose of primary antibody was then added to each well. To assay HA reactivity, 100 µl/well of a 1:10,000 dilution of non-immune or PR/8/34 convalescent immune serum in ADB was added. PR/8/34 immune serum was generated in BALB/c mice (Jackson Laboratory, Bar Harbor, ME) that received an experimentally determined sublethal challenge dose of $8x10^1$ egg infectious dosages (EID) of PR/8/34 influenza virus. Animals were then allowed to convalesce for >21days post-infection at which time immune serum was isolated and clarified.

For ELISA of flagellin or the 6x histidine tag , monoclonal antibody against 6x His (Invitrogen; Carlsbad, CA), or flagellin (Inotek; Beverly, MA) was added at 1 µg/ml in ADB (100 µl/well) and the plates were incubated for 1 hr at room temperature or overnight at 4°C. The plates were then washed three times with PBS-T. HRP-labeled goat anti-mouse IgG antibodies (Jackson Immunochemical; West Grove, PA) diluted in ADB were added (100 µl/well) and the plates were incubated at room temperature for 1 hour. The plates were washed three times with PBS-T. After adding TMB Ultra substrate (Pierce; Rockford, IL) and monitoring color development, A_{450} was measured on a microplate spectrophotometer (FARCyte, GE/Amersham; Piscataway, NJ).

Results and Discussion

Expression and purification of recombinant HA proteins: The results of *Baculovirus* expression of recombinant STF2.HA fusion proteins are summarized in Table 9. All proteins were expressed in moderate to high yield. Purity after metal chelate chromatography was generally good and endotoxin levels were far below the acceptable limit of 0.1 EU/µg. All STF2 fusion proteins demonstrated potent *in vitro* TLR5 activity.

Table 9

Protein	SEQ ID NO:	Yield (mg)	Purity est. (%)	Endotoxin (EU/μg)	TLR activity (EC_{50}, ng/ml)
STF2.HA1-1His(PR8)	151	12	>80	0.01	3
ngSTF2.HA1-1His(PR8)	152	24	>90	<0.01	30
STF2.HA1-2His(PR8)	153	13.2	>90	<0.01	9
STF2.HA1-2*mut*His(PR8)	155	25	>95	<0.01	900
ngSTF2.HA1-2*mut*His(PR8)	156	6	>90	0.01	9

STF2.HA proteins for *Baculovirus* expression were engineered with and without N-linked glycosylation sites in order to determine the effect of this posttranslational modification on protein expression, folding, and biological activity. ngSTF2.HA1-2.*mut*His(PR8) (SEQ ID NO: 156), which is non-glycosylated, appears to be expressed at a lower level in *Baculovirus* supernatants than the corresponding glycosylated protein, suggesting that glycosylation may influence the folding or secretion of this protein. However, the glycosylated and non-glycosylated versions of this protein appear to react equally with convalescent antiserum from influenza PR8-infected mice when analyzed by western blot. Furthermore, ngSTF2.HA1-1His(PR8) (SEQ ID NO: 152) expressed at higher levels than its glycosylated counterpart, indicating that any effect of glycosylation on expression of these proteins may not be easily generalized.

In addition to the proteins expressed and purified at large scale, small-scale *Baculovirus* P1 supernatants of STF2.HA1-3His(PR8) (SEQ ID NO: 157) and STF2.HA1-3*mut*His(PR8) (SEQ ID NO: 158) were analyzed for HA antibody reactivity.. As seen with the STF2.HA1-3(PR8) protein (SEQ ID NO: 92) expressed in *E. coli* (See Example #4), these proteins showed very poor reactivity with PR8 convalescent mouse serum on western blots of conditioned medium. This result further confirms that STF2.HA1-3 proteins, whether expressed in a prokaryotic or eukaryotic host, are unable to fold properly and display native epitopes.

Antigenicity: Several STF2.HA proteins produced in *Baculovirus* and *E. coli* were analyzed by ELISA. All STF2 fusion proteins reacted equally well with anti-flagellin antibody (Inotek; Beverly, MA), demonstrating equal protein loading in the assay. Proper folding and display of antibody epitopes in STF2.HA fusion proteins was then examined by measuring reactivity of each protein with antisera from mice which had recovered from PR/8/34 influenza infection. Reactivity with PR/8/34 convalescent mouse serum was equivalent for *Baculovirus*-expressed STF2.HA1-1His(PR8) (SEQ ID NO: 151), ngSTF2.HA1-1His(PR8) (SEQ ID NO: 152), and ngSTF2HA1-2*mut*His(PR8) (SEQ ID NO: 156), and *E. coli*-expressed STF2.HA1-2(PR8) (SEQ ID NO: 90) (see Exmple #4). Thus, serum antibody recognition of the HA head domain may be focused on a region encompassed by the HA1-2 construct. Further, *in vitro* refolding of STF2.HA1-2 (SEQ ID NO: 90) produced in *E. coli* results in a protein with equivalent immune reactivity (and therefore similar folded conformation) to a protein processed and secreted from a eukaryotic host (*Baculovirus*). In addition, serum antibody recognition of the HA head domain may not appear to be dependent on glycosylation and, at least for PR/8/34 HA, may not be negatively affected by glycosylation. HAs (hemaglutinins) from different influenza strains have variable patterns of N-linked glycosylation. (glycosylation may negatively affect immune recognition of other HAs. Serum antibody recognition of HA may be unaffected by N-glycosylation. It is possible glycosylation may otherwise influence the effectiveness, such as the half-life of a vaccine.

EXAMPLE 9: IMMUNOGENICITY AND EFFICACY OF RECOMBINANT FLAGELLIN-HEMAGGLUTININ FUSION PROTEINS PRODUCED IN *BACULOVIRUS*

Material and Methods

Animal studies: Female BALB/c mice (Jackson Laboratory, bar Harbor, ME) were used at the age of 6-8 weeks. Mice were divided into groups of 10 and received inguinal subcutaneous immunizations on days 0 and 14 as follows:

1) PBS (phosphate buffered saline).
2) 3 μg of STF2.HA1-2(PR8) *E. coli* (SEQ ID NO: 90) in PBS
3) 3 μg of STF2.HA1-1(PR8) *Baculovirus* (SEQ ID NO: 151) in PBS
4) 3 μg of ngSTF2.HA1-1(PR8) *Baculovirus* (SEQ ID NO: 152) in PBS
5) 0.3 μg of ngSTF2.HA1-1(PR8) *Baculovirus* (SEQ ID NO: 152) in PBS
6) 3 μg of STF2.HA1-1(PR8) *E. coli* (SEQ ID NO: 89) in PBS
7) 0.3 μg of STF2.HA1-1(PR8) *E. coli* (SEQ ID NO: 89) in PBS

Mice were bled on days 10 (primary) and 21 (boost), and sera were clarified by clotting and centrifugation and stored at -20°C. Immunized animals were challenged on day 28 with an LD_{90} ($8x10^3$ EID) of PR/8/34 virus administered intranasally (See Example 6). Mice were monitored daily for 21 days following the challenge for survival, weight loss and clinical presentation.

Serum antibody determination: HA-specific IgG levels were determined by ELISA. 96-well ELISA plates (Costar (Cat#9018), Corning, NY) were coated overnight at 4°C with 100 μl/well HA0sHis protein produced in *Drosophila* (SEQ ID NO 176) in PBS (5 μg/ml). Plates were blocked with 200 μl/well of Assay Diluent Buffer (ADB; BD Pharmingen (Cat# 555213), San Diego, CA) for one hour at room temperature. The plates were washed three times in PBS containing 0.05% Tween-20 (PBS-T). Dilutions of the sera in ADB were added (100 μl/well) and the plates were incubated overnight at 4°C. The plates were washed three times with PBS-T. HRP-labeled goat anti-mouse IgG antibodies (Jackson Immunochemical (Cat# 115-035-146)) diluted in ADB were added (100 μl/well) and the plates were incubated at room temperature for 1 hour. The plates were washed three times with PBS-T. After adding TMB Ultra substrate (Pierce (Cat# 34028), Rockford, PA) and monitoring color development, A_{450} was measured on a Tecan Farcyte (Durham, NC) microplate spectrophotometer.

Results and Discussion

Immunogenicity and efficacy of STF2.HA1-1(PR8) expressed in *E. coli* and *Baculovirus*: BALB/c mice were immunized s.c. with 3.0 or 0.3 μg of indicated recombinant fusion proteins on days 0 and 14. On day 21 mice were bled and HA-specific IgG titers were examined by ELISA against HA0sHis (SEQ ID NO 176) expressed in *Drosophila*. Figures 12A,12B shows the proteins induced varying levels of HA-specific IgG with ngSTF2.HA1-1(PR8) (SEQ ID NO: 152) inducing the strongest response similar to that observed in animals immunized with STF2.HA1-2(PR8) expressed in *E. coli* (SEQ ID NO: 90). Interestingly, animals immunized with STF2.HA1-1(PR8) containing the intact glycosylation sequences expressed in *Baculovirus* (SEQ ID NO: 151) elicited little to no detectable antibody responses to HA or flagellin, in marked contrast to that observed in animals similarly immunized with STF2.HA1-1(PR8) where the consensus glycosylation sequences had been removed (SEQ ID NO: 152).

The immunized mice were challenged on day 28 with an LD_{90} ($8x10^3$ EID) of PR/8/34 virus administered intranasally. Mice were monitored daily for 21 days following the challenge for survival, weight loss and clinical presentation. As shown in Figures 13A, 13B, 13C, PBS-immunized mice showed signs of infection (weight loss and lower clinical scores) as early as three days after challenge and died by day 21, while mice immunized with 3.0 μg of recombinant STF2.HA1-2(PR8) (SEQ ID NO: 90) fusion protein demonstrated enhanced protection with little to no observable clinical signs of infection or weight loss. Although animals immunized with STF2.HA1-1 fusion proteins demonstrated significantly higher efficacy than animals immunized with PBS alone, these animals demonstrated lower efficacy in terms of survival, clinical scores and overall weight loss, compared to animals receiving STF2.HA1-2(PR8) (SEQ ID NO: 90). These results demonstrate that HA subunits linked to STF2 in *Baculovirus* are immunogenic and efficacious, and suggest that STF2.HA1-1 (SEQ ID NO: 89) expressed in *E. coli* or *Baculovirus* is less immunogenic and efficacious than STF2.HA1-2 expressed in *E. coli*(SEQ ID NO: 90) *or Baculovirus*.

EXAMPLE 10: PURIFICATION OF STF2.HA1-2 (IND) (SEQ ID NO: 159)

Materials and Methods

Cell banking: *E. coli* cells engineered to express STF2.HA1-2(IND) (SEQ ID NO: 159) were adapted to culture in proprietary MRSF media, banked as glycerol stocks in 1 mL aliquots, and stored at -70°C.

MRSF media, pH 7.0

Composition	g/L
Glucose	10
KH_2PO_4	7.8
$(NH_4)_2SO4$	2.33
Citric Acid	1.0
$MgSO_4(7H_20)$	1.0
$CaCl_2$	0.04
Trace Metals	1ml
Thiamine HCl	0.01
Kanamycin	0.0075

Trace Metal Solution 1000x

Component	g/L
EDTA, disodium	5
$FeSO_4(7H_2O)$	10
$ZnSO_4(7H_2O)$	2
$MnSO_4(H_2O)$	2
$CoCl_2(6H_2O)$	0.2
$CuSO_4(5H_2O)$	0.1
$Na_2MoO_4(2H_2O)$	0.2
H_3BO_3	0.1
H_2O	1000ml

Cell scale-up: Two vials (2 ml, banked in MRSF media) of *E. coli* cells engineered to express STF2.HA1-2 (IND) (SEQ ID NO: 159) were retrieved from --70°C and inoculated into 500 ml of MRSF medium. The cells were expanded by incubation at 37°C for 11.5 hours, at which time the OD_{600} was 3.2. The cell scale up provided biomass to inoculate the production bioreactor used in the fermentation.

Fermentation: A 250 mL aliquot of scaled up cells was used to inoculate a 12 L working volume bioreactor containing 10 L of proprietary MRBR media. The culture was incubated with constant stirring while maintaining pH at 7.0±0.1, temperature at 30±0.1°C, and dissolved oxygen (DO) at no less than 30%, and was

run in batch mode until the glucose was exhausted. Glucose exhaustion is indicated by a sudden decrease in O_2 demand, and confirmed by measurement of glucose concentration on a YSI 2700 Select glucose meter. The OD_{600} for the culture at this time was 20.1 AU.

Thirty minutes after glucose exhaustion, proprietary feed media was pumped into the bioreactor at a controlled rate until 2 L of feed was added. The incubation continued for 5.5 hours under glucose-limiting conditions. At an OD_{600} of 34 AU, protein expression in the culture was induced by the addition of IPTG to a final concentration of 2 mM and incubation with constant stirring for 2 hours. At the conclusion of the induction period, cells were harvested by centrifugation in an Avanti JP-20 XP centrifuge for 20 minutes at 10,000 g. The final OD_{600} for the culture was 58 AU and the total protein concentration as determined by BCA was 6.73 mg/mL. 1.16 kg of cell paste was recovered from 10.8 L harvest and frozen at -20°C. The total bioreactor run time was 17 hours. STF2.HA1-2 (IND) (SEQ ID NO: 159) production was confirmed by SDS PAGE.

MRBR Media, pH 7.0

Composition	g/L
Glucose	20
KH_2PO_4	2.2
$(NH_4)_2SO_4$	4.5
Citric Acid	1.0
$MgSO_4(7H_20)$	1.0
$CaCl_2$	0.04
Trace Metals	1ml
Thiamine HCl	0.01
Antifoam	0.05
Kanamycin	0.0075

Feed media, pH 6.0 – additions to MRBR media (2 L final volume)

Composition	g/L
Glucose	160
KH_2PO_4	5.6
DL-Alanine	50.0
Citric Acid	3.0
$MgSO_4(7H_20)$	1.0 (added as 78g/L solution)
$CaCl_2$	2.5 (added as 80g/L solution)
Trace Metals	24ml
$FeSO_4$ $7H_2O$	0.75

Cell disruption and clarification: The cell paste was thawed and resuspended into a 15% suspension of 50 mM Tris, 25 mM NaCl, pH 8. The slurry was homogenized in an APV1000 homogenizer three times at 12,000 psi. The lysate was adjusted to pH 4.0 by the addition of acetic acid and centrifuged to separate soluble from insoluble material. STF2.HA1-2 (IND) (SEQ ID NO: 159) along with most other proteins partition to the insoluble fraction. The soluble material was discarded and the insoluble material (pellet) was dissolved in two times the initial lysate volume of 50 mM acetate, 1% TritonX-100, 8 M Urea, pH 4. After mixing in the homogenizer at 0 psi, the sample was centrifuged and the supernatant was collected, filtered, and stored at 4°C. The presence of STF2.HA1-2 (IND) (SEQ ID NO: 159) in the proper phases (precipitate or supernatant) was confirmed by SDS PAGE.

Cation exchange chromatography (CEX): The clarified material was applied to a 50 ml Poros 50HS column. This step was designed to reduce the endotoxin concentration, reduce the nucleic acid concentration and increase the STF2.HA1-2 (IND) (SEQ ID NO: 159) purity of the eluate as compare to the load. After loading, the column was washed with five column volumes of 50 mM acetate, 1% Triton X-100, 8 M Urea, pH 4 followed by 10 column volumes of 50 mM acetate, 8 M Urea, pH 4. The protein was eluted in a 0-100% gradient to 1 M NaCl. The target protein eluted at approximately 300 mM NaCl. The major peak (eluate) was pooled and stored at 4°C. The presence of STF2.HA1-2 (IND) (SEQ ID NO: 159) in the elution was confirmed by SDS PAGE. STF2.HA1-2 (IND) (SEQ ID NO: 159) purification

across this step was determined by comparing the number and intensity of protein bands in the load material to those in the elution.

Endotoxin removal in cation exchange chromatography (CEX): Endotoxin is removed from eluting STF2.HA1-2 (IND) (SEQ ID NO: 159) with the 50 mM acetate, 1% Triton X-100, 8 M Urea, pH 4 wash and with proper elution conditions. This was tested using STF2.HA1-2 (VN) (SEQ ID NO: 95) Comparing two separate cation exchange runs (Toso SP650M resin) with and without the 50 mM acetate, 1% Triton X-100, 8 M Urea, pH 4 wash (ER wash vs. no ER wash), the endotoxin concentration in the ER wash elution at 150 mM NaCl shows significantly less endotoxin concentration that the 250 mm NaCl elution from the same run, or in either of the elutions without the 50 mM acetate, 1% Triton X-100, 8 M Urea, pH 4 wash.

Run	Elution (mM NaCl)	[Endotoxin] (EU/mL)	[Endotoxin] (EU/mg)	[protein] by BCA (g/L)	UV 280/260
Load	N/A	4.4×10^6	1.8×10^6	2.39	-
No ER wash	100	7.9×10^6	4.2×10^6	1.44	1.05
	250	8.0×10^6	4.1×10^6	1.12	2.76
ER Wash	150	5,112	2,888	1.77	0.982
	250	24,440	N/A	N/A	0.545

Refolding: The CEX eluate was dialyzed against 50 mM Tris, 8 M Urea, pH 8 buffer, and refolded by rapid dilution into 10 volumes of 50 mM Tris, pH 8.

Anion Exchange Chromatography (AEX): The refolded protein was applied to a 25 ml GE Source Q column. This step is designed to separate properly folded STF2.HA1-2 (IND) (SEQ ID NO: 159) from aggregates and host cell proteins and to reduce the endotoxin concentration of the eluate as compared to the load. After loading, the column was washed with five column volumes of 50 mM Tris, 25 mM NaCl, pH 8. The protein was eluted in a 0-100% gradient to 1 M NaCl. A peak enriched for properly folded STF2.HA1-2(IND) (SEQ ID NO: 159) with low endotoxin (1.7 EU/mL) eluted at approximately 150 mM NaCl. A second peak eluting at approximately 170 to 270 mM NaCl contains other proteins including STF2.HA1-2 (IND) (SEQ ID NO: 159) aggregates. Fractions in first peak were pooled at stored at 4°C. Fractions from the first peak were kept for further processing. The presence of STF2.HA1-2 (IND) (SEQ ID NO: 159) in the first and

second elutions was confirmed by SDS PAGE. STF2.HA1-2 (IND) (SEQ ID NO: 159) purification across this step was determined by comparing the number and intensity of protein bands in the load material to those in the elution.

Preparative size exclusion chromatography (SEC): The AEX eluate was run on a 320 ml SEC column in 2 x 5 mL fractions against TBS. This step is designed to separate correctly folded STF2.HA1-2 (IND) (SEQ ID NO: 159) from other proteins. Each run resulted in elution of a single major peak which was collected, pooled, and filtered through a 0.22µm filter. The product was stored at -70°C.

The presence of STF2.HA1-2 (IND) (SEQ ID NO: 159) in the elution was confirmed by SDS PAGE. STF2.HA1-2 (IND) (SEQ ID NO: 159) purification across this step was determined by comparing the number and intensity of protein bands in the load material to those in the elution. The final material profile on SDS PAGE is a strong single band.

Results and Discussion

Final characterization of protein: Recombinant STF2.HA1-2(IND) (SEQ ID NO: 159) protein was expressed in *E. coli* and purified to homogeneity. The final product was formulated in 1x TBS (27.7 mM Tris, 2.7 mM KCl, 137 mM NaCl, pH 7.4; made from 10x stock, Teknova catalog # T9530) and stored at -70°C in aliquots of 1.0 ml. The final yield was 15.75 ml (45 ml at a concentration of 0.35 mg/ml), with an endotoxin level of 1.1 EU/mg protein as determined by the LAL method. The protein retained TLR5 biological activity as measured in a cell-based assay of cytokine release

SDS-PAGE: Protein identity was determined, and purity estimated, by SDS-PAGE. An aliquot sample was diluted in SDS-PAGE sample buffer and diluent. The samples were boiled for 5 minutes and loaded onto a 4 to 12% SDS polyacrylamide gel (Invitrogen NuPage) and electrophoresed. The gel was stained with Coomassie R-250 to visualize protein bands. Positive results for STF2.HA1-2 (IND) (SEQ ID NO: 159) were indicated by a band at about 75 kDa as compared to BioRad precision plus standards run on the same gel. Purity is estimated by visually comparing the intensity of the STF2.HA1-2 (IND) (SEQ ID NO: 159) band to all other bands.

TLR5 bioactivity assay: HEK293 cells constitutively express TLR5, and secrete several soluble factors, including IL-8, in response to TLR5 signaling. Cells were seeded in 96-well microplates (50,000 cells/well), and test proteins were added and incubated overnight. The next day, the conditioned medium was harvested, transferred to a clean 96-well microplate, and frozen at -20°C. After thawing, the conditioned medium was assayed for the presence of IL-8 in a sandwich ELISA using an anti-human IL-8 matched antibody pair (Pierce, #M801E and #M802B) following the manufacturer's instructions. Optical density was measured using a microplate spectrophotometer (FARCyte, Amersham).

Endotoxin assay: Endotoxin levels were determined using the QCL-1000 Quantitative Chromogenic LAL test kit (Cambrex), following the manufacturer's instructions for the microplate method.

Protein assay: Total protein concentration was determined using the Micro BCA (bicinchoninic acid) Assay (Pierce) in the microplate format, using bovine serum albumin as a standard, according to the manufacturer's instructions.

UV280/260 assay: This assay is a measure of the nucleic acid concentration in a sample as compared to the protein concentration. The absorbance maximum occurs at approximately 260 nm for nucleic acids, while the absorbance maximum occurs for proteins at approximately 280 nm. A higher ratio between the absorbance at 280 nm versus the absorbance at 260 nm indicates a sample enriched in protein over nucleic acids.

EXAMPLE 11: CLONING, EXPRESSION, AND BIOCHEMICAL CHARACTERIZATION OF RECOMBINANT FLAGELLIN-HEMAGGLUTININ FUSION PROTEINS PRODUCED IN *DROSOPHILA*

Materials and Methods

Cloning and expression of HA in *Drosophila*: RNA was isolated from the influenza A strain A/Puerto Rico/8/34 using QIAamp MinElute Virus Spin Kit (Qiagen, Cat#57704) following the manufacturer's instructions. The HA RNA was reverse transcribed using SuperScript III First Strand Synthesis System for RT-PCR (Invitrogen; Cat #18080-051) or SuperScript III One-Step RT-PCR System with Platinum-Taq High Fidelity (Invitrogen; Cat# 12574-030) following the manufacturer's instructions, and the cDNA was cloned using the TOPO cloning

vector PCRII Zero Blunt (Invitrogen). The HA0s fragment (HA gene without the signal sequence and trans-membrane domain) was then sub-cloned into pMT/STF2Δ at both the C-terminus and NH2 terminus of the cassette using PCR-based strategy to generate the constructs STF2Δ.HA0s (PR8) (SEQ ID NO: 160) and HA0s(PR8).STF2Δ (SEQ ID NO: 167) respectively. The HA gene was also cloned into the pMT/Bip/V5-his vector without STF2Δ to generate the construct HA0s.his (SEQ ID NO: 170). Table 10 list the constructs and the primers used in their construction.

Table 10: PR8 HA constructs for expression in *Drosophila*

SEQ ID NO:	Construct	FOR Primer SEQ ID NO:	REV Primer SEQ ID NO:	DNA Template SEQ ID NO:
160	STF2Δ.HA0sHis	161	162	163
164	STF2Δ.HA0s	165	166	163
167	HA0s.STF2Δ	168	169	163
170	HA0sHis	171	172	163
173	HA0s	174	175	163

Recombinant protein expression from *Drosophila* Dmel-2 cells: The HA0s.His (SEQ ID NO: 170) and STF2Δ.HA0s (SEQ ID NO: 164) plasmids were co-transfected with the pCoBlast plasmid into *Drosophila* Dmel-2 cells to generate Dmel-2 HA0sHis(PR8) and Dmel-2 STF2Δ.HA0s(PR8) stable cells. Dmel-2 stable cells were expanded from adherent to shaker cultures in selection media [*Drosophila* Sfm medium (Invitrogen; Carlsbad, CA) + 25 μg/ml blasticidin] and then expanded to a 12 L production scale.

Recombinant protein expression was induced by the addition of $CuSO_4$ to a final concentration of 0.5 mM. After incubation for 72 hours cells were removed from the conditioned medium by centrifugation. The conditioned medium was then clarified by passing through a 0.22 μm filter and stored at 4°C.

Purification of HA0sHis(PR8) (SEQ ID NO: 176): Conditioned medium was applied to a Chelating Sepharose column (GE/Amersham Biosciences; Piscataway,

NJ) charged with $NiSO_4$ and equilibrated in Buffer A [20 mM Tris, pH 8.0, 0.5 M NaCl] and eluted in a linear gradient with Buffer B [Buffer A + 0.5 M imidazole]. The protein was further purified by fractionation on a Superdex 200 SEC column (GE/Amersham Biosciences; Piscataway, NJ) equilibrated in 1X Tris-buffered saline (TBS), pH 8.0. HA0s.His_6(PR8) fractions were pooled, aliquoted and stored at -80°C

SDS-PAGE and western blot analysis: Protein identity was determined, and purity estimated, by SDS-PAGE. An aliquot of 5 µg of each sample was diluted in SDS-PAGE sample buffer with or without 100 mM DTT as a reductant. The samples were boiled for 5 minutes and loaded onto a 10% SDS polyacrylamide gel (LifeGels; French's Forrest, New South Wales, AUS) and electrophoresed. The gel was stained with Coomassie R-250 (Bio-Rad; Hercules, CA) to visualize protein bands. For western blot, 0.5 µg/lane total protein was electrophoresed as described above and the gels were then electro-transfered to a PVDF membrane and blocked with 5% (w/v) dry milk before probing with anti-flagellin antibody (Inotek; Beverly, MA), anti-His_6.antibody (Invitrogen; Carlsbad, CA) or influenza A PR/8/34 convalescent immune serum (described below under Protein Antigenicity ELISA). After probing with alkaline phosophatase-conjugated secondary antibodies (Pierce; Rockford, IL), protein bands were visualized with an alkaline phosphatase chromogenic substrate (Promega; Madison, WI).

Protein assay: Total protein concentration was determined using the Micro BCA (bicinchoninic acid) Assay (Pierce; Rockford, IL) in the microplate format, using bovine serum albumin as a standard, according to the manufacturer's instructions.

Endotoxin assay: Endotoxin levels were determined using the QCL-1000 Quantitative Chromogenic LAL test kit (Cambrex; E. Rutherford, NJ), following the manufacturer's instructions for the microplate method.

TLR5 bioactivity assay: HEK293 cells constitutively express TLR5, and secrete several soluble factors, including IL-8, in response to TLR5 signaling. Cells were seeded in 96-well microplates (50,000 cells/well), and recombinant *drosophila* conditioned medium containing either HA0sHis(PR8) (SEQ ID NO: 176) or STF2Δ.HA0s(PR8) (SEQ ID NO: 177) was added. . The next day, the conditioned medium was harvested, transferred to a clean 96-well microplate, and frozen at -

20°C. After thawing, the conditioned medium was assayed for the presence of IL-8 in a sandwich ELISA using an anti-human IL-8 matched antibody pair (Pierce; Rockford, IL, #M801E and #M802B) following the manufacturer's instructions. Optical density was measured using a microplate spectrophotometer (FARCyte, GE/Amersham; Piscataway, NJ).

Protein antigenicity ELISA: Purified HA0s.His_6(PR8) (SEQ ID NO: 176) was tested by ELISA to determine if the recombinant protein displayed correctly folded epitopes of HA. 96-well ELISA plates were coated overnight at 4°C with serial dilutions in PBS (100 µl/well) of HA0s.His_6(PR8) (SEQ ID NO: 176) protein starting at 5 µg/ml. Plates were blocked with 200 µl/well of Assay Diluent Buffer (ADB; BD Pharmingen) for one hour at room temperature, then washed three times in PBS-T. A fixed dose of primary antibody was then added to each well. To assay HA reactivity, 100 µl/well of a 1:10,000 dilution of non-immune or PR/8/34 convalescent immune serum in ADB was added. PR/8/34 immune serum was generated in BALB/c mice (Jackson Laboratory, Bar Harbor, ME) that received an experimentally determined sublethal challenge dose of $8x10^1$ egg infectious dosages (EID) of PR/8/34 influenza virus.

Animals were then allowed to convalesce for >21days post-infection at which time immune serum was isolated and clarified. For ELISA of the 6x histidine tag , monoclonal antibody against 6x His (Invitrogen; Carlsbad, CA), or flagellin (Inotek; Beverly, MA) was added at 1 µg/ml in ADB (100 µl/well) and the plates were incubated for 1 hr at room temperature or overnight at 4°C. The plates were then washed three times with PBS-T. HRP-labeled goat anti-mouse IgG antibodies (Jackson Immunochemical; West Grove, PA) diluted in ADB were added (100 µl/well) and the plates were incubated at room temperature for 1 hour. The plates were washed three times with PBS-T. After adding TMB Ultra substrate (Pierce; Rockford, IL) and monitoring color development, A_{450} was measured on a microplate spectrophotometer (FARCyte, GE/Amersham; Piscataway, NJ).

Results and Discussion

Characterization of *Drosphila*-expressed HA0s proteins: Western blot analysis of conditioned media with anti-His_6 antibody (Invitrogen; Carlsbad, CA)

confirmed expression of HA0sHis(PR8) (SEQ ID NO: 176) and western blot with anti-flagellin antibody (Inotek; Beverly, MA) confirmed expression of STF2Δ.HA0s(PR8) (SEQ ID NO: 177) by *Drosophila* Dmel-2 cells transfected with the corresponding expression plasmid. Both proteins were recognized by western blot in non-reduced form with mouse PR/8/34 convalescent immune serum while reduction of the proteins with DTT abrogated recognition. This result indicates correct disulfide bonding of the two secreted proteins. The STF2Δ.HA0s(PR8) (SEQ ID NO: 177) conditioned medium showed significant *in vitro* TLR5 activity while the HA0sHis$_6$(PR8) (SEQ ID NO: 176) medium did not, as expected. Finally, purified HA0sHis$_6$(PR8) (SEQ ID NO: 176) protein showed significant reactivity with influenza A PR8/34 convalescent immune serum by ELISA. These results indicate that both HA0s proteins are secreted from *Drosophila* Dmel-2 cells in a properly folded form.

EXAMPLE 12: IMMUNOGENICITY AND EFFICACY OF RECOMBINANT FLAGELLIN-HEMAGGLUTININ FUSION PROTEINS REPRESENTING VIRAL STRAIN A/VIET NAM/1203/04

Materials and Methods

Animal studies: Female BALB/c mice were used at the age of 6-8 weeks. Mice were divided into groups of 15 and received subcutaneous (s.c) immunizations on days 0 and 14 as follows:

- TBS (phosphate buffered saline)
- No immunization
- 10 μg of STF2.HA1-2(VN) (SEQ ID NO: 95) in TBS
- 3.0 μg of STF2.HA1-2(VN) (SEQ ID NO: 95) in PBS
- 1 μg of STF2.HA1-2(VN) (SEQ ID NO: 95) in TBS

Mice were bled on day 21 and sera were clarified by clotting and centrifugation and stored at -20°C.

Serum antibody determination: HA-specific IgG levels were determined by ELISA. 96-well ELISA plates (Costar (Cat #9018) Corning, NY) were coated overnight at 4°C with 100 μl/well recombinant HA protein from A/Vietnam/1203/04, produced in *Baculovirus*, (BEIR catalog number NR-660) in

PBS (1 μg/ml). Plates were blocked with 300 μl/well of Assay Diluent Buffer (ADB; BD Pharmingen, (Cat#: 555213)(San Diego, CA) for two hours at 25°C. The plates were washed three times in PBS+ 0.05% (v/v) Tween 20 (PBS-T). Dilutions of the sera in ADB were added (100 μl/well) and the plates were incubated for 1.5 hours at 25°C. The plates were washed three times with PBS-T. HRP-labeled goat anti-mouse IgG antibodies (Jackson Immunochemical, West Grove, PA (Cat#: 115-035-146)) diluted in ADB were added (100 μl/well) and the plates were incubated at 25°C for 30 minutes. The plates were washed three times with PBS-T. After adding TMB Ultra substrate (Pierce (Cat 34028), Rockford, IL)) and monitoring color development, A_{450} was measured on a SpectraMax 190 (Molecular Devices, Sunnyvale, CA) microplate spectrophotometer.

Influenza virus challenge of mice: To assess efficacy, mice were challenged on day 28 by intranasal administration of 10 LD_{90} (10 x dose lethal to 90% of mice) ($6x10^3$ EID) of influenza A/Viet Nam/1203/04. Animals were monitored daily for 21 days following the challenge for survival and clinical presentation.

Preparation of Influenza H5N1 Stocks: Influenza A/Vietnam/1203/04 (H5N1) was obtained from the Centers for Disease Control and Prevention (Atlanta, GA). Stocks of the virus were prepared in 10-day-old embryonated chicken eggs, and aliquots of 0.5 ml were stored at -80°C until use.

Determination of Viral Titers: $TCID_{50}$ (50% tissue culture infectious dose) was determined utilizing MDCK cells seeded into the wells of 96-well plate and grown to confluency. A series of log_{10} dilutions of the virus was inoculated onto the plate in quadruplicates. The plate was then incubated for 48-72 hours. The $TCID_{50}$ was determined by identifying the dilution at which quadruplicate wells were ½ positive and ½ negative for viral growth. The EID_{50} was determined by inoculating log_{10} dilutions of the stock into 2-4 eggs per dilution. Eggs were incubated for 40-48 hours, and 1 ml of the allantoic fluid was harvested from each egg. EID_{50} was calculated as the dilution at which ½ of the eggs were negative and ½ of the eggs were positive for infectious virus.

Results and Discussion

Induction of HA-specific IgG response following immunization with STF2.HA1-2(VN) (SEQ ID NO: 95): The immunogenicity of STF2.HA1-2(VN) (SEQ ID NO: 95) was examined by immunizing BALB/c mice (15/group) subcutaneously on day 0 and 14 with a dose range of 10, 3, and 1 μg protein. Negative control groups of mice were immunized with TBS or not immunized. HA-specific IgG responses were examined 7 days post boost (Day 21) by ELISA. The results demonstrate that immunization with 10, 3 or 1 μg of STF2.HA1-2(VN) (SEQ ID NO: 95) induced consistent and significant HA-specific IgG responses in a dose-dependent manner (Figures 14-16).

Immunization with STF2.HA1-2(VN) (SEQ ID NO: 95) provides protection from a lethal challenge with influenza A: The serological analysis described above demonstrated that immunization of mice with STF2.HA1-2(VN) (SEQ ID NO: 95) generated an antibody response that recognized native HA. In order to evaluate efficacy, the same mice were challenged on day 28 with 10 LD_{90} ($6x10^3$EID) of A/Viet Nam/1203/04 virus administered intra-nasally. Mice were monitored daily for 21 days following the challenge for survival and clinical presentation. As shown in Figure 17, TBS-immunized or naïve mice showed signs of infection (weight loss and lower clinical scores) as early as five days post-challenge and all mice died by day nine post-challenge. For the mice immunized with STF2.HA1-2(VN), there is a clear relationship between the dose received and efficacy. Mice immunized with 10 μg of STF2.HA1-2(VN) (SEQ ID NO: 95) demonstrated markedly enhanced protection. Mice immunized with 3 or 1 μg of STF2.HA1-2(VN) (SEQ ID NO: 95) demonstrated a modest to a negligible impact on efficacy, respectively. Clinical parameters were similarly related to dose and animals receiving the highest dose of STF2.HA1-2(VN) (SEQ ID NO: 95) exhibited the mildest signs of disease. These results demonstrate that *E. coli*-expressed STF2.HA1-2(VN) (SEQ ID NO: 95) induces HA-specific immune responses that successfully protect BALB/c mice from a lethal challenge with virulent influenza A virus *in vivo*.

SEQ ID NO: 1 A/Puerto Rico/8/34 HA

MKANLLVLLSALAAADADTICIGYHANNSTDTVDTVLEKNVTVTHSVNLLEDSHNGKLCRL
KGIAPLQLGKCNIAGWLLGNPECDPLLPVRSWSYIVETPNSENGICYPGDFIDYEELREQL
SSVSSFERFEIFPKESSWPNHNTNGVTAACSHEGKSSFYRNLLWLTEKEGSYPKLKNSYVN
KKGKEVLVLWGIHHPPNSKEQQNLYQNENAYVSVVTSNYNRRFTPEIAERPKVRDQAGRMN
YYWTLLKPGDTIIFEANGNLIAPMYAFALSRGFGSGIITSNASMHECNTKCQTPLGAINSS
LPYQNIHPVTIGECPKYVRSAKLRMVTGLRNIPSIQSRGLFGAIAGFIEGGWTGMIDGWYG
YHHQNEQGSGYAADQKSTQNAINGITNKVNTVIEKMNIQFTAVGKEFNKLEKRMENLNKKV
DDGFLDIWTYNAELLVLLENERTLDFHDSNVKNLYEKVKSQLKNNAKEIGNGCFEFYHKCD
NECMESVRNGTYDYPKYSEESKLNREKVDGVKLESMGIYQILAIYSTVASSLVLLVSLGAI
SFWMCSNGSLQCRICI

SEQ ID NO: 2 A/Viet Nam/1203/2004 HA

MEKIVLLFAIVSLVKSDQICIGYHANNSTEQVDTIMEKNVTVTHAQDILEKKHNGKLCDLD
GVKPLILRDCSVAGWLLGNPMCDEFINVPEWSYIVEKANPVNDLCYPGDFNDYEELKHLLS
RINHFEKIQIIPKSSWSSHEASLGVSSACPYQGKSSFFRNVVWLIKKNSTYPTIKRSYNNT
NQEDLLVLWGIHHPNDAAEQTKLYQNPTTYISVGTSTLNQRLVPRIATRSKVNGQSGRMEF
FWTILKPNDAINFESNGNFIAPEYAYKIVKKGDSTIMKSELEYGNCNTKCQTPMGAINSSM
PFHNIHPLTIGECPKYVKSNRLVLATGLRNSPQRERRRKKRGLFGAIAGFIEGGWQGMVDG
WYGYHHSNEQGSGYAADKESTQKAIDGVTNKVNSIIDKMNTQFEAVGREFNNLERRIENLN
KKMEDGFLDVWTYNAELLVLMENERTLDFHDSNVKNLYDKVRLQLRDNAKELGNGCFEFYH
KCDNECMESVRNGTDYPQYSEEARLKREEISGVKLESIGIYQILSIYSTVASSLALAIMVA
GLSLWMCSNGSLQCR

SEQ ID NO: 3 A/Indonesia/5/2005 HA

MEKIVLLLAIVSLVKSDQICIGYHANNSTEQVDTIMEKNVTVTHAQDILEKTHNGKLCDLD
GVKPLILRDCSVAGWLLGNPMCDEFINVPEWSYIVEKANPTNDLCYPGSFNDYEELKHLLS
RINHFEKIQIIPKSSWSDHEASSGVSSACPYLGSPSFFRNVVWLIKKNSTYPTIKKSYNNT
NQEDLLVLWGIHHPNDAAEQTRLYQNPTTYISIGTSTLNQRLVPKIATRSKVNGQSGRMEF
FWTILKPNDAINFESNGNFIAPEYAYKIVKKGDSAIMKSELEYGNCNTKCQTPMGAINSSM
PFHNIHPLTIGECPKYVKSNRLVLATGLRNSPQRESRRKKRGLFGAIAGFIEGGWQGMVDG
WYGYHHSNEQGSGYAADKESTQKAIDGVTNKVNSIIDKMNTQFEAVGREFNNLERRIENLN
KKMEDGFLDVWTYNAELLVLMENERTLDFHDSNVKNLYDKVRLQLRDNAKELGNGCFEFYH
KCDNECMESIRNGTYNYPQYSEEARLKREEISGVKLESIGTYQILSIYSTVASSLALAIMM
AGLSLWMCSNGSLQCRICI

SEQ ID NO: 4 A/NewCaledonia/20/1999

MKAKLLVLLCTFTATYADTICIGYHANNSTDTVDTVLEKNVTVTHSVNLLEDSHNGKLCLL
KGIAPLQLGNCSVAGWILGNPECELLISKESWSYIVETPNPENGTCYPGYFADYEELREQL
SSVSSFERFEIFPKESSWPNHTVTGVSASCSHNGKSSFYRNLLWLTGKNGLYPNLSKSYVN
NKEKEVLVLWGVHHPPNIGNQRALYHTENAYVSVVSSHYSRRFTPEIAKRPKVRDQEGRIN
YYWTLLEPGDTIIFEANGNLIAPWYAFALSRGFGSGIITSNAPMDECDAKCQTPQGAINSS
LPFQNVHPVTIGECPKYVRSAKLRMVTGLRNIPSIQSRGLFGAIAGFIEGGWTGMVDGWYG
YHHQNEQGSGYAADQKSTQNAINGITNKVNSVIEKMNTQFTAVGKEFNKLERRMENLNKKV
DDGFLDIWTYNAELLVLLENERTLDFHDSNVKNLYEKVKSQLKNNAKEIGNGCFEFYHKCN
NECMESVKNGTYDYPKYSEESKLNREKIDGVKLESMGVYQILAIYSTVASSLVLLVSLGAI
SFWMCSNGSLQCRICI

SEQ ID NO: 5 A/South Carolina/1/18HA

MEARLLVLLCAFAATNADTICIGYHANNSTDTVDTVLEKNVTVTHSVNLLEDSHNGKLCKL
KGIAPLQLGKCNIAGWLLGNPECDLLLTASSWSYIVETSNSENGTCYPGDFIDYEELREQL
SSVSSFEKFEIFPKTSSWPNHETTKGVTAACSYAGASSFYRNLLWLTKKGSSYPKLSKSYV
NNKGKEVLVLWGVHHPPTGTDQQSLYQNADAYVSVGSSKYNRRFTPEIAARPKVRDQAGRM

NYYWTLLEPGDTITFEATGNLIAPWYAFALNRGSGSGIITSDAPVHDCNTKCQTPHGAINS
SLPFQNIHPVTIGECPKYVRSTKLRMATGLRNIPSIQSRGLFGAIAGFIEGGWTGMIDGWY
GYHHQNEQGSGYAADQKSTQNAIDGITNKVNSVIEKMNTQFTAVGKEFNNLERRIENLNKK
VDDGFLDIWTYNAELLVLLENERTLDFHDSNVRNLYEKVKSQLKNNAKEIGNGCFEFYHKC
DDACMESVRNGTYDYPKYSEESKLNREEIDGVKLESMGVYQILAIYSTVASSLVLLVSLGA
ISFWMCSNGSLQCRICI

SEQ ID NO: 6 A/Wisconsin/67/2005 HA

QKLPGNDNSTATLCLGHHAVPNGTIVKTITNDQIEVTNATELVQSSSTGGICDSPHQILDG
ENCTLIDALLGDPQCDGFQNKKWDLFVERSKAYSNCYPYDVPDYASLRSLVASSGTLEFND
ESFNWTGVTQNGTSSACKRRSNNSFFSRLNWLTHLKFKYPALNVTMPNNEKFDKLYIWGVH
HPGTDNDQIFLHAQASGRITVSTKRSQQTVIPNIGSRPRIRNIPSRISIYWTIVKPGDILL
INSTGNLIAPRGYFKIRSGKSSIMRSDAPIGKCNSECITPNGSIPNDKPFQNVNRITYGAC
PRYVKQNTLKLATGMRNVPEKQTR

SEQ ID NO: 7 A/X31 subtype H3N2 HA

QDLPGNDNSTATLCLGHHAVPNGTLVKTITDDQIEVTNATELVQSSSTGKICNNPHRILDG
IDCTLIDALLGDPHCDVFQNETWDLFVERSKAFSNCYPYDVPDYASLRSLVASSGTLEFIT
EGFTWTGVIQNGGSNACKRGPGSGFFSRLNWLTKSGSTYPVLNVTMPNNDNFDKLYIWGIH
HPSTNQEQTSLYVQASGRVTVSTRRSQQTIIPNIGSRPWVRGLSSRISIYWTIVKPGDVLV
INSNGNLIAPRGYFKMRTGKSSIMRSDAPIDTCISECITPNGSIPNDKPFQNVNKITYGAC
PKYVKQNTLKLATGMRNVPEKQT

SEQ ID NO: 8 PR/8 HA1-1

SHNGKLCRLKGIAPLQLGKCNIAGWLLGNPECDPLLPVRSWSYIVETPNSENGICYPGDFI
DYEELREQLSSVSSFERFEIFPKESSWPNHNTNGVTAACSHEGKSSFYRNLLWLTEKEGSY
PKLKNSYVNKKGKEVLVLWGIHHPPNSKEQQNLYQNENAYVSVVTSNYNRRFTPEIAERPK
VRDQAGRMNYYWTLLKPGDTIIFEANGNLIAPMYAFALSRGFGSGIITSNASMHECNTKCQ
TPLGAINSSLPYQNIHPVTIGECPKYVR

SEQ ID NO: 9 PR/8 HA1-2

KGIAPLQLGKCNIAGWLLGNPECDPLLPVRSWSYIVETPNSENGICYPGDFIDYEELREQL
SSVSSFERFEIFPKESSWPNHNTNGVTAACSHEGKSSFYRNLLWLTEKEGSYPKLKNSYVN
KKGKEVLVLWGIHHPPNSKEQQNLYQNENAYVSVVTSNYNRRFTPEIAERPKVRDQAGRMN
YYWTLLKPGDTIIFEANGNLIAPMYAFALSRGFGSGIITS

SEQ ID NO: 10 PR/8 HA1-3

NSENGICYPGDFIDYEELREQLSSVSSFERFEIFPKESSWPNHNTNGVTAACSHEGKSSFY
RNLLWLTEKEGSYPKLKNSYVNKKGKEVLVLWGIHHPPNSKEQQNLYQNENAYVSVVTSNY
NRRFTPEIAERPKVRDQAGRMNYYWTLLKPGDTIIFEANGNLIAPMYAFALSRG

SEQ ID NO: 11 VN HA1-1

EKKHNGKLCDLDGVKPLILRDCSVAGWLLGNPMCDEFINVPEWSYIVEKANPVNDLCYPGD
FNDYEELKHLLSRINHFEKIQIIPKSSWSSHEASLGVSSACPYQGKSSFFRNVVWLIKKNS
TYPTIKRSYNNTNQEDLLVLWGIHHPNDAAEQTKLYQNPTTYISVGTSTLNQRLVPRIATR
SKVNGQSGRMEFFWTILKPNDAINFESNGNFIAPEYAYKIVKKGDSTIMKSELEYGNCNTK
CQTPMGAINSSMPFHNIHPLTIGECPKYVK

SEQ ID NO: 12 VN HA1-2

GVKPLILRDCSVAGWLLGNPMCDEFINVPEWSYIVEKANPVNDLCYPGDFNDYEELKHLLS
RINHFEKIQIIPKSSWSSHEASLGVSSACPYQGKSSFFRNVVWLIKKNSTYPTIKRSYNNT
NQEDLLVLWGIHHPNDAAEQTKLYQNPTTYISVGTSTLNQRLVPRIATRSKVNGQSGRMEF
FWTILKPNDAINFESNGNFIAPEYAYKIVKKGDSTIMKSE

SEQ ID NO: 13 VN HA1-3
NDLCYPGDFNDYEELKHLLSRINHFEKIQIIPKSSWSSHEASLGVSSACPYQGKSSFFRNV
VWLIKKNSTYPTIKRSYNNTNQEDLLVLWGIHHPNDAAEQTKLYQNPTTYISVGTSTLNQR
LVPRIATRSKVNGQSGRMEFFWTILKPNDAINFESNGNFIAPEYAYKIVKKG

SEQ ID NO: 14 IND HA1-1
EKTHNGKLCDLDGVKPLILRDCSVAGWLLGNPMCDEFINVPEWSYIVEKANPTNDLCYPGS
FNDYEELKHLLSRINHFEKIQIIPKSSWSDHEASSGVSSACPYLGSPSFFRNVVWLIKKNS
TYPTIKKSYNNTNQEDLLVLWGIHHPNDAAEQTRLYQNPTTYISIGTSTLNQRLVPKIATR
SKVNGQSGRMEFFWTILKPNDAINFESNGNFIAPEYAYKIVKKGDSAIMKSELEYGNCNTK
CQTPMGAINSSMPFHNIHPLTIGECPKYVK

SEQ ID NO: 15 IND HA1-2
GVKPLILRDCSVAGWLLGNPMCDEFINVPEWSYIVEKANPTNDLCYPGSFNDYEELKHLLS
RINHFEKIQIIPKSSWSDHEASSGVSSACPYLGSPSFFRNVVWLIKKNSTYPTIKKSYNNT
NQEDLLVLWGIHHPNDAAEQTRLYQNPTTYISIGTSTLNQRLVPKIATRSKVNGQSGRMEF
FWTILKPNDAINFESNGNFIAPEYAYKIVKKGDSAIMKSE

SEQ ID NO: 16 IND HA1-3
NDLCYPGSFNDYEELKHLLSRINHFEKIQIIPKSSWSDHEASSGVSSACPYLGSPSFFRNV
VWLIKKNSTYPTIKKSYNNTNQEDLLVLWGIHHPNDAAEQTRLYQNPTTYISIGTSTLNQR
LVPKIATRSKVNGQSGRMEFFWTILKPNDAINFESNGNFIAPEYAYKIVKKG

SEQ ID NO: 17 NC HA1-1
SHNGKLCLLKGIAPLQLGNCSVAGWILGNPECELLISKESWSYIVETPNPENGTCYPGYFA
DYEELREQLSSVSSFERFEIFPKESSWPNHTVTGVSASCSHNGKSSFYRNLLWLTGKNGLY
PNLSKSYVNNKEKEVLVLWGVHHPPNIGNQRALYHTENAYVSVVSSHYSRRFTPEIAKRPK
VRDQEGRINYYWTLLEPGDTIIFEANGNLIAPWYAFALSRGFGSGIITSNAPMDECDAKCQ
TPQGAINSSLPFQNVHPVTIGECPKYVR

SEQ ID NO: 18 NC HA1-2
KGIAPLQLGNCSVAGWILGNPECELLISKESWSYIVETPNPENGTCYPGYFADYEELREQL
SSVSSFERFEIFPKESSWPNHTVTGVSASCSHNGKSSFYRNLLWLTGKNGLYPNLSKSYVN
NKEKEVLVLWGVHHPPNIGNQRALYHTENAYVSVVSSHYSRRFTPEIAKRPKVRDQEGRIN
YYWTLLEPGDTIIFEANGNLIAPWYAFALSRGFGSGIITS

SEQ ID NO: 19 NC HA1-3
NPENGTCYPGYFADYEELREQLSSVSSFERFEIFPKESSWPNHTVTGVSASCSHNGKSSFY
RNLLWLTGKNGLYPNLSKSYVNNKEKEVLVLWGVHHPPNIGNQRALYHTENAYVSVVSSHY
SRRFTPEIAKRPKVRDQEGRINYYWTLLEPGDTIIFEANGNLIAPWYAFALSRG

SEQ ID NO: 20 WIS HA1-1
QSSSTGGICDSPHQILDGENCTLIDALLGDPQCDGFQNKKWDLFVERSKAYSNCYPYDVPD
YASLRSLVASSGTLEFNDESFNWTGVTQNGTSSACKRRSNNSFFSRLNWLTHLKFKYPALN
VTMPNNEKFDKLYIWGVHHPGTDNDQIFLHAQASGRITVSTKRSQQTVIPNIGSRPRIRNI
PSRISIYWTIVKPGDILLINSTGNLIAPRGYFKIRSGKSSIMRSDAPIGKCNSECITPNGS
IPNDKPFQNVNRITYGACPRYVK

SEQ ID NO: 21 WIS HA1-2

SPHQILDGENCTLIDALLGDPQCDGFQNKKWDLFVERSKAYSNCYPYDVPDYASLRSLVAS
SGTLEFNDESFNWTGVTQNGTSSACKRRSNNSFFSRLNWLTHLKFKYPALNVTMPNNEKFD
KLYIWGVHHPGTDNDQIFLHAQASGRITVSTKRSQQTVIPNIGSRPRIRNIPSRISIYWTI
VKPGDILLINSTGNLIAPRGYFKIRSGKSSIMRSD

SEQ ID NO: 22 WIS HA1-3

SNCYPYDVPDYASLRSLVASSGTLEFNDESFNWTGVTQNGTSSACKRRSNNSFFSRLNWLT
HLKFKYPALNVTMPNNEKFDKLYIWGVHHPGTDNDQIFLHAQASGRITVSTKRSQQTVIPN
IGSRPRIRNIPSRISIYWTIVKPGDILLINSTGNLIAPRGYFKIRSG

SEQ ID NO: 23 A/Puerto Rico/8/34 HA0s

DTICIGYHANNSTDTVDTVLEKNVTVTHSVNLLEDSHNGKLCRLKGIAPLQLGKCNIAGWL
LGNPECDPLLPVRSWSYIVETPNSENGICYPGDFIDYEELREQLSSVSSFERFEIFPKESS
WPNHNTNGVTAACSHEGKSSFYRNLLWLTEKEGSYPKLKNSYVNKKGKEVLVLWGIHHPPN
SKEQQNLYQNENAYVSVVTSNYNRRFTPEIAERPKVRDQAGRMNYYWTLLKPGDTIIFEAN
GNLIAPMYAFALSRGFGSGIITSNASMHECNTKCQTPLGAINSSLPYQNIHPVTIGECPKY
VRSAKLRMVTGLRNIPSIQSRGLFGAIAGFIEGGWTGMIDGWYGYHHQNEQGSGYAADQKS
TQNAINGITNKVNTVIEKMNIQFTAVGKEFNKLEKRMENLNKKVDDGFLDIWTYNAELLVL
LENERTLDFHDSNVKNLYEKVKSQLKNNAKEIGNGCFEFYHKCDNECMESVRNGTYDYPKY
SEESKLNREKVDGVKLESMGIYQ

SEQ ID NO: 24 PR/8 HA1-2*mut*

KGAAPLQLGKCNIAGWLLGNPECDPLLPVRSWSDIAETPNSENGICYPGDFIDYEELREQL
SSVSSFERFEIFPKESSWPNHNTNGVTAACSHEGKSSFYRNLLWLTEKEGSYPKLKNSYVN
KKGKEVLVLWGIHHPPNSKEQQNLYQNENAYVSVVTSNYNRRFTPEIAERPKVRDQAGRMN
YYWTLLKPGDTIIFEANGNLIAPMYAFALSRGFGSGIITS

SEQ ID NO: 25 PR/8 HA1-3*mut*

NSENEICYPGDFIDKEELREQLSSVSSFERFEIFPKESSWPNHNTNGVTAACSHEGKSSFY
RNLLWLTEKEGSYPKLKNSYVNKKGKEVLVLWGIHHPPNSKEQQNLYQNENAYVSVVTSNY
NRRFTPEIAERPKVRDQAGRMNYYWTLLKPGDTIIFEANGNLIAPMYAAALSRG

SEQ ID NO: 26 VN HA1-2*mut*

GAKPLSLRDCSVAGWLLGNPMCDEFINVPEWSDIAEKANPVNDLCYPGDFNDYEELKHLLS
RINHFEKIQIIPKSSWSSHEASLGVSSACPYQGKSSFFRNVVWLIKKNSTYPTIKRSYNNT
NQEDLLVLWGIHHPNDAAEQTKLYQNPTTYISVGTSTLNQRLVPRIATRSKVNGQSGRMEF
FWTILKPNDAINFESNGNFIAPEYAYKIVKKGDSTIMKSE

SEQ ID NO: 27 VN HA1-3*mut*

NDLCYPGDFNDKEELKHLLSRINHFEKIQIIPKSSWSSHEASLGVSSACPYQGKSSFFRNV
VWLIKKNSTYPTIKRSYNNTNQEDLLVLWGIHHPNDAAEQTKLYQNPTTYISVGTSTLNQR
LVPRIATRSKVNGQSGRMEFFWTILKPNDAINFESNGNFIAPEYAYKIAKKG

SEQ ID NO: 28 IND HA1-2*mut*

GAKPLSLRDCSVAGWLLGNPMCDEFINVPEWSDIAEKANPTNDLCYPGSFNDYEELKHLLS
RINHFEKIQIIPKSSWSDHEASSGVSSACPYLGSPSFFRNVVWLIKKNSTYPTIKKSYNNT
NQEDLLVLWGIHHPNDAAEQTRLYQNPTTYISIGTSTLNQRLVPKIATRSKVNGQSGRMEF
FWTILKPNDAINFESNGNFIAPEYAYKIVKKGDSAIMKSE

SEQ ID NO: 29 IND HA1-3*mut*

NDLCYPGSFNDKEELKHLLSRINHFEKIQIIPKSSWSDHEASSGVSSACPYLGSPSFFRNV
VWLIKKNSTYPTIKKSYNNTNQEDLLVLWGIHHPNDAAEQTRLYQNPTTYISIGTSTLNQR
LVPKIATRSKVNGQSGRMEFFWTILKPNDAINFESNGNFIAPEYAYKIAKKG

SEQ ID NO: 30 NC HA1-2*mut*

KGAAPLQLGNCSVAGWILGNPECELLISKESWSDIAETPNPENGTCYPGYFADYEELREQL
SSVSSFERFEIFPKESSWPNHTVTGVSASCSHNGKSSFYRNLLWLTGKNGLYPNLSKSYVN
NKEKEVLVLWGVHHPPNIGNQRALYHTENAYVSVVSSHYSRRFTPEIAKRPKVRDQEGRIN
YYWTLLEPGDTIIFEANGNLIAPWYAFALSRGFGSGIITS

SEQ ID NO: 31 NC HA1-3*mut*

NPENETCYPGYFADKEELREQLSSVSSFERFEIFPKESSWPNHTVTGVSASCSHNGKSSFY
RNLLWLTGKNGLYPNLSKSYVNNKEKEVLVLWGVHHPPNIGNQRALYHTENAYVSVVSSHY
SRRFTPEIAKRPKVRDQEGRINYYWTLLEPGDTIIFEANGNLIAPWYAAALSRG

SEQ ID NO: 32 WIS HA1-2*mut*

SPHQALDGENCTLIDALLGDPQCDGFQNKKWDDFAERSKAYSNCYPYDVPDYASLRSLVAS
SGTLEFNDESFNWTGVTQNGTSSACKRRSNNSFFSRLNWLTHLKFKYPALNVTMPNNEKFD
KLYIWGVHHPGTDNDQIFLHAQASGRITVSTKRSQQTVIPNIGSRPRIRNIPSRISIYWTI
VKPGDILLINSTGNLIAPRGYFKIRSGKSSIMRSD

SEQ ID NO: 33 WIS HA1-3*mut*

SNCYPKDSPDEASLRSLVASSGTDEFNDESFNWTGVTQNGTSSACKRRSNNSFFSRLNWLT
HLKFKYPALNVTMPNNEKFDKLYIWGVHHPGTDNDQIFLHAQASGRITVSTKRSQQTVIPN
IGSRPRIRNIPSRISIYWTIVKPGDILLINSTGNLIAPRGYFKIRSG

SEQ ID NO: 34 A/Swine/Hong Kong/9/98 HA (ACCESSION BAB85618)

DKICIGYQSTNSTETVDTLTETNVPVTHAKELLHTEHNGMLCATNLGHPLILDTCTIEGLI
YGNPSCDLLLGGREWSYIVERPSAVNGMCYPGNVENLEELRSLFSSASSYQRIQIFPDTIW
NVSYSGTSKACSDSFYRSMRWLTQKNNAYPIQDAQYTNNRGKSILFMWGINHPPTDTVQTN
LYTRTDTTTSVTTEDINRTFKPVIGPRPLVNGLHGRIDYYWSVLKPGQTLRVRSNGNLIAP
WYGHILSGESHGRILKTDLNSGNCVVQCQTERGGLNTTLPFHNVSKYAFGNCPKYVGVKSL
KLAVGLRNVPARSSR

SEQ ID NO: 35 A/Aichi/2/1968 HA (ACCESSION BAF37221)

MKTIIALSYIFCLALGQDLPGNDNSTATLCLGHHAVPNGTLVKTITDDQIEVTNATELVQS
SSTGKICNNPHRILDGIDCTLIDALLGDPHCDVFQNETWDLFVERSKAFSNCYPYDVPDYA
SLRSLVASSGTLEFITEGFTWTGVTQNGGSNACKRGPGSGFFSRLNWLTKSGSTYPVLNVT
MPNNDNFDKLYIWGIHHPSTNQEQTSLYVQASGRVTVSTRRSQQTIIPNIGSRPWVRGLSS
RISIYWTIVKPGDVLVINSNGNLIAPRGYFKMRTGKSSIMRSDAPIDTCISECITPNGSIP
NDKPFQNVNKITYGACPKYVKQNTLKLATGMRNVPEKQTRGLFGAIAGFIENGWEGMIDGW
YGFRHQNSEGTGQAADLKSTQAAIDQINGKLNRVIEKTNEKFHQIEKEFSEVEGRIQDLEK
YVEDTKIDLWSYNAELLVALENQHTIDLTDSEMNKLFEKTRRQLRENAEEMGNGCFKIYHK
CDNACIESIRNGTYDHDVYRDEALNNRFQIKGVELKSGYKDWILWISFAISCFLLCVVLLG
FIMWACQRGNIRCNICI

SEQ ID NO: 36 B/Lee/40 HA (Accession NP_056660)

MKAIIVLLMVVTSNADRICTGITSSNSPHVVKTATQGEVNVTGVIPLTTTPTKSHFANLKG
TQTRGKLCPNCFNCTDLDVALGRPKCMGNTPSAKVSILHEVKPATSGCFPIMHDRTKIRQL
PNLLRGYENIRLSTSNVINTETAPGGPYKVGTSGSCPNVANGNGFFNTMAWVIPKDNNKTA
INPVTVEVPYICSEGEDQITVWGFHSDDKTQMERLYGDSNPQKFTSSANGVTTHYVSQIGG

FPNQTEDEGLKQSGRIVVDYMVQKPGKTGTIVYQRGILLPQKVWCASGRSKVIKGSLPLIG
EADCLHEKYGGLNKSKPYYTGEHAKAIGNCPIWVKTPLKLANGTKYRPPAKLLKERGFFGA
IAGFLEGGWEGMIAGWHGYTSHGAHGVAVAADLKSTQEAINKITKNLNYLSELEVKNLQRL
SGAMNELHDEILELDEKVDDLRADTISSQIELAVLLSNEGIINSEDEHLLALERKLKKMLG
PSAVEIGNGCFETKHKCNQTCLDRIAAGTFNAGDFSLPTFDSLNITAASLNDDGLDNHTIL
LYYSTAASSLAVTLMIAIFIVYMVSRDNVSCSICL

SEQ ID NO: 37 C/Johannesburg/1/66 (ACCESSION AAW73083)

MFFSLLLVLGLTEAEKIKICLQKQVNSSFSLHNGFGGNLYATEEKRMFELVKPKAGASVLN
QSTWIGFGDSRTDKSNSAFPRSADVSAKTADKFRSLSGGSLMLSMFGPPGKVDYLYQGCGK
HKVFYEGVNWSPHAAINCYRKNWTDIKLNFQKNIYELASQSHCMSLVNALDKTIPLQVTAG
TAGNCNNSFLKNPALYTQEVKPSENKCGKENLAFFTLPTQFGTYECKLHLVASCYFIYDSK
EVYNKRGCDNYFQVIYDSSGKVVGGLDNRVSPYTGNSGDTPTMQCDMLQLKPGRYSVRSSP
RFLLMPERSYCFDMKEKGPVTAVQSIWGKGRESDYAVDQACLSTPGCMLIQKQKPYIGEAD
DHHGDQEMRELLSGLDYEARCISQSGWVNETSPFTEKYLLPPKFGRCPLAAKEESIPKIPD
GLLIPTSGTDTTVTKPKSRIFGIDDLIIGLLFVAIVEAGIGGYLLGSRKESGGGVTKESAE
KGFEKIGNDIQILKSSINIAIEKLNDRISHDEQAIRDLTLEIENARSEALLGELGIIRALL
VGNISIGLQESLWELASEITNRAGDLAVEVSPGCWIIDNNICDQSCQNFIFKFNETAPVPT
IPPLDTKIDLQSDPFYWGSSLGLAITATISLAALVISGIAICRTK

SEQ ID NO: 38 B/Lee/40 HA1-1

TTTPTKSHFANLKGTQTRGKLCPNCFNCTDLDVALGRPKCMGNTPSAKVSILHEVKPATSG
CFPIMHDRTKIRQLPNLLRGYENIRLSTSNVINTETAPGGPYKVGTSGSCPNVANGNGFFN
TMAWVIPKDNNKTAINPVTVEVPYICSEGEDQITVWGFHSDDKTQMERLYGDSNPQKFTSS
ANGVTTHYVSQIGGFPNQTEDEGLKQSGRIVVDYMVQKPGKTGTIVYQRGILLPQKVWCAS
GRSKVIKGSLPLIGEADCLHEKYGGLNKSKPYYTGEHAKAIGNCPIWVK

SEQ ID NO: 39 B/Lee/40 HA1-2

KGTQTRGKLCPNCFNCTDLDVALGRPKCMGNTPSAKVSILHEVKPATSGCFPIMHDRTKIR
QLPNLLRGYENIRLSTSNVINTETAPGGPYKVGTSGSCPNVANGNGFFNTMAWVIPKDNNK
TAINPVTVEVPYICSEGEDQITVWGFHSDDKTQMERLYGDSNPQKFTSSANGVTTHYVSQI
GGFPNQTEDEGLKQSGRIVVDYMVQKPGKTGTIVYQRGILLPQKVWCASGRSKVIKG

SEQ ID NO: 40 A/Aichi/2/68 HA1-1

QSSSTGKICNNPHRILDGIDCTLIDALLGDPHCDVFQNETWDLFVERSKAFSNCYPYDVPD
YASLRSLVASSGTLEFITEGFTWTGVTQNGGSNACKRGPGSGFFSRLNWLTKSGSTYPVLN
VTMPNNDNFDKLYIWGIHHPSTNQEQTSLYVQASGRVTVSTRRSQQTIIPNIGSRPWVRGL
SSRISIYWTIVKPGDVLVINSNGNLIAPRGYFKMRTGKSSIMRSDAPIDTCISECITPNGS
IPNDKPFQNVNKITYGACPKYVK

SEQ ID NO: 41 A/Aichi/2/68 HA1-2

HRILDGIDCTLIDALLGDPHCDVFQNETWDLFVERSKAFSNCYPYDVPDYASLRSLVASSG
TLEFITEGFTWTGVTQNGGSNACKRGPGSGFFSRLNWLTKSGSTYPVLNVTMPNNDNFDKL
YIWGIHHPSTNQEQTSLYVQASGRVTVSTRRSQQTIIPNIGSRPWVRGLSSRISIYWTIVK
PGDVLVINSNGNLIAPRGYFKMRTGKSSIMRSD

SEQ ID NO: 42 B/Malaysia/2506/2004 HA (Accession ISDN126672)

IVLLMVVTSNADRICTGITSSNSPHVVKTATQGEVNVTGVIPLTTTPTKSHFANLKGTETR
GKLCPKCLNCTDLDVALGRPKCTGNIPSARVSILHEVRPVTSGCFPIMHDRTKIRQLPNLL
RGYEHIRLSTHNVINAENAPGGSYKIGTSGSCPNVTNGNGFFATMAWAVPKNDNNKTATNS
LTIEVPYICTEGEDQITVWGFHSDNEAQMAKLYGDSKPQKFTSSANGVTTHYVSQIGGFPN

QTEDGGLPQSGRIVVDYMVQKSGKTGTITYQRGILLPQKVWCASGRSKVIKGSLPLIGEAD
CLHEKYGGLNKSKPYYTGEHAKAIGNCPIWVKTPLKLANGTKYRPPAKLLKER

SEQ ID NO: 43 B/Ohio/1/2005 HA (Accession ISDN133312)

DRICTGITSSNSPHVVKTATQGEVNVTGVIPLTTTPTKSHFANLKGTKTRGKLCPKCLNCT
DLDVALGRPKCTGNIPSAEVSILHEVRPVTSGCFPIMHDRTKIRQLPNLLRGYEHIRLSTH
NVINAEKAPGGPYKIGTSGSCPNVTNGNGFFATMAWAVPKNDNNKTATNSLTIEVPYICTE
GEDQITIWGFHSDSETQMAKLYGDSKPQKFTSSANGVTTHYVSQIGGFPNQTEDGGLPQSG
RIVVDYMVQKSGKTGTITYQRGILLPQKVWCASGRSKVIKGSLPLIGEADCLHEKYGGLNK
SKPYYTGEHAKAIGNCPIWVKTPLKLANGTKYRPPAKLLKERGF

SEQ ID NO: 44 B/Shanghai/361/2002 HA (Accession ISDN80784)

DRICTGITSSNSPHVVKTATQGEVNVTGVIPLTTTPIKSHFANLKGTRTRGKLCPDCLNCT
DLDVALGRPMCVGTTPSAKASILHEVRPVTSGCFPIMHDRTKIRQLPNLLRGYENIRLSTQ
NVIDAEKALGGPYRLGTSGSCPNATSKSGFFATMAWAVPKDNNKNATNPLTVEVPYICTEG
EDQITVWGFHSDDKTQMKNLYGDSNPQKFTSSANGVTTHYVSQIGGFPDQTEDGGLPQSGR
IVVDYMVQKPGKTGTIVYQRGVLLPQKVWCASGRSKVIKGSLPLIGEADCLHEKYGGLNKS
KPYYTGEHAKAIGHCPIWVKTPLKLANGTKYRP

SEQ ID NO: 45 B/Malaysia/2506/2004 HA1-1

TTTPTKSHFANLKGTETRGKLCPKCLNCTDLDVALGRPKCTGNIPSARVSILHEVRPVTSG
CFPIMHDRTKIRQLPNLLRGYEHIRLSTHNVINAENAPGGSYKIGTSGSCPNVTNGNGFFA
TMAWAVPKNDNNKTATNSLTIEVPYICTEGEDQITVWGFHSDNEAQMAKLYGDSKPQKFTS
SANGVTTHYVSQIGGFPNQTEDGGLPQSGRIVVDYMVQKSGKTGTITYQRGILLPQKVWCA
SGRSKVIKGSLPLIGEADCLHEKYGGLNKSKPYYTGEHAKAIGNCPIWVK

SEQ ID NO: 46 B/Malaysia/2506/2004 HA1-2

KGTETRGKLCPKCLNCTDLDVALGRPKCTGNIPSARVSILHEVRPVTSGCFPIMHDRTKIR
QLPNLLRGYEHIRLSTHNVINAENAPGGSYKIGTSGSCPNVTNGNGFFATMAWAVPKNDNN
KTATNSLTIEVPYICTEGEDQITVWGFHSDNEAQMAKLYGDSKPQKFTSSANGVTTHYVSQ
IGGFPNQTEDGGLPQSGRIVVDYMVQKSGKTGTITYQRGILLPQKVWCASGRSKVIKG

SEQ ID NO: 47 B/Ohio/1/2005 HA1-1

TTTPTKSHFANLKGTKTRGKLCPKCLNCTDLDVALGRPKCTGNIPSAEVSILHEVRPVTSG
CFPIMHDRTKIRQLPNLLRGYEHIRLSTHNVINAEKAPGGPYKIGTSGSCPNVTNGNGFFA
TMAWAVPKNDNNKTATNSLTIEVPYICTEGEDQITIWGFHSDSETQMAKLYGDSKPQKFTS
SANGVTTHYVSQIGGFPNQTEDGGLPQSGRIVVDYMVQKSGKTGTITYQRGILLPQKVWCA
SGRSKVIKGSLPLIGEADCLHEKYGGLNKSKPYYTGEHAKAIGNCPIWVK

SEQ ID NO: 48 B/Ohio/1/2005 HA1-2

KGTKTRGKLCPKCLNCTDLDVALGRPKCTGNIPSAEVSILHEVRPVTSGCFPIMHDRTKIR
QLPNLLRGYEHIRLSTHNVINAEKAPGGPYKIGTSGSCPNVTNGNGFFATMAWAVPKNDNN
KTATNSLTIEVPYICTEGEDQITIWGFHSDSETQMAKLYGDSKPQKFTSSANGVTTHYVSQ
IGGFPNQTEDGGLPQSGRIVVDYMVQKSGKTGTITYQRGILLPQKVWCASGRSKVIKG

SEQ ID NO: 49 B/Shanghai/361/2002 HA1-1

TTTPIKSHFANLKGTRTRGKLCPDCLNCTDLDVALGRPMCVGTTPSAKASILHEVRPVTSG
CFPIMHDRTKIRQLPNLLRGYENIRLSTQNVIDAEKALGGPYRLGTSGSCPNATSKSGFFA
TMAWAVPKDNNKNATNPLTVEVPYICTEGEDQITVWGFHSDDKTQMKNLYGDSNPQKFTSS
ANGVTTHYVSQIGGFPDQTEDGGLPQSGRIVVDYMVQKPGKTGTIVYQRGVLLPQKVWCAS
GRSKVIKGSLPLIGEADCLHEKYGGLNKSKPYYTGEHAKAIGHCPIWVK

SEQ ID NO: 50 B/Shanghai/361/2002 HA1-2

KGTRTRGKLCPDCLNCTDLDVALGRPMCVGTTPSAKASILHEVRPVTSGCFPIMHDRTKIR
QLPNLLRGYENIRLSTQNVIDAEKALGGPYRLGTSGSCPNATSKSGFFATMAWAVPKDNNK
NATNPLTVEVPYICTEGEDQITVWGFHSDDKTQMKNLYGDSNPQKFTSSANGVTTHYVSQI
GGFPDQTEDGGLPQSGRIVVDYMVQKPGKTGTIVYQRGVLLPQKVWCASGRSKVIKG

SEQ ID NO: 51 STF2.blp

ATGGCACAAGTAATCAACACTAACAGTCTGTCGCTGCTGACCCAGAATAACCTGAACAAA
TCCCAGTCCGCACTGGGCACCGCTATCGAGCGTCTGTCTTCTGGTCTGCGTATCAACAGC
GCGAAAGACGATGCGGCAGGTCAGGCGATTGCTAACCGTTTCACCGCGAACATCAAAGGT
CTGACTCAGGCTTCCCGTAACGCTAACGACGGTATCTCCATTGCGCAGACCACTGAAGGC
GCGCTGAACGAAATCAACAACAACCTGCAGCGTGTGCGTGAACTGGCGGTTCAGTCTGCT
AACAGCACCAACTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATCACCCAGCGCCTG
AACGAAATCGACCGTGTATCCGGCCAGACTCAGTTCAACGGCGTGAAAGTCCTGGCGCAG
GACAACACCCTGACCATCCAGGTTGGCGCCAACGACGGTGAAACTATCGATATCGATCTG
AAGCAGATCAACTCTCAGACCCTGGGTCTGGACTCACTGAACGTGCAGAAAGCGTATGAT
GTGAAAGATACAGCAGTAACAACGAAAGCTTATGCCAATAATGGTACTACACTGGATGTA
TCGGGTCTTGATGATGCAGCTATTAAAGCGGCTACGGGTGGTACGAATGGTACGGCTTCT
GTAACCGGTGGTGCGGTTAAATTTGACGCAGATAATAACAAGTACTTTGTTACTATTGGT
GGCTTTACTGGTGCTGATGCCGCCAAAAATGGCGATTATGAAGTTAACGTTGCTACTGAC
GGTACAGTAACCCTTGCGGCTGGCGCAACTAAAACCACAATGCCTGCTGGTGCGACAACT
AAAACAGAAGTACAGGAGTTAAAAGATACACCGGCAGTTGTTTCAGCAGATGCTAAAAAT
GCCTTAATTGCTGGCGGCGTTGACGCTACCGATGCTAATGGCGCTGAGTTGGTCAAAATG
TCTTATACCGATAAAAATGGTAAGACAATTGAAGGCGGTTATGCGCTTAAAGCTGGCGAT
AAGTATTACGCCGCAGATTACGATGAAGCGACAGGAGCAATTAAAGCTAAAACTACAAGT
TATACTGCTGCTGACGGCACTACCAAAACAGCGGCTAACCAACTGGGTGGCGTAGACGGT
AAAACCGAAGTCGTTACTATCGACGGTAAAACCTACAATGCCAGCAAAGCCGCTGGTCAT
GATTTCAAAGCACAACCAGAGCTGGCGGAAGCAGCCGCTAAAACCACCGAAAACCCGCTG
CAGAAAATTGATGCCGCGCTGGCGCAGGTGGATGCGCTGCGCTCTGATCTGGGTGCGGTA
CAAAACCGTTTCAACTCTGCTATCACCAACCTGGGCAATACCGTAAACAATCTGTCTGAA
GCGCGTAGCCGTATCGAAGATTCCGACTACGCGACCGAAGTTTCCAACATGTCTCGCGCG
CAGATTCTGCAGCAGGCCGGTACTTCCGTTCTGGCGCAGGCTAACCAGGTCCCGCAGAAC
GTGCTGAGCCTGTTACGT

SEQ ID NO: 52 STF2.SG

ATGGCACAAGTAATCAACACTAACAGTCTGTCGCTGCTGACCCAGAATAACCTGAACAAA
TCCCAGTCCGCACTGGGCACCGCTATCGAGCGTCTGTCTTCTGGTCTGCGTATCAACAGC
GCGAAAGACGATGCGGCAGGTCAGGCGATTGCTAACCGTTTCACCGCGAACATCAAAGGT
CTGACTCAGGCTTCCCGTAACGCTAACGACGGTATCTCCATTGCGCAGACCACTGAAGGC
GCGCTGAACGAAATCAACAACAACCTGCAGCGTGTGCGTGAACTGGCGGTTCAGTCTGCT
AACAGCACCAACTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATCACCCAGCGCCTG
AACGAAATCGACCGTGTATCCGGCCAGACTCAGTTCAACGGCGTGAAAGTCCTGGCGCAG
GACAACACCCTGACCATCCAGGTTGGCGCCAACGACGGTGAAACTATCGATATCGATCTG
AAGCAGATCAACTCTCAGACCCTGGGTCTGGACTCACTGAACGTGCAGAAAGCGTATGAT
GTGAAAGATACAGCAGTAACAACGAAAGCTTATGCCAATAATGGTACTACACTGGATGTA
TCGGGTCTTGATGATGCAGCTATTAAAGCGGCTACGGGTGGTACGAATGGTACGGCTTCT
GTAACCGGTGGTGCGGTTAAATTTGACGCAGATAATAACAAGTACTTTGTTACTATTGGT
GGCTTTACTGGTGCTGATGCCGCCAAAAATGGCGATTATGAAGTTAACGTTGCTACTGAC
GGTACAGTAACCCTTGCGGCTGGCGCAACTAAAACCACAATGCCTGCTGGTGCGACAACT
AAAACAGAAGTACAGGAGTTAAAAGATACACCGGCAGTTGTTTCAGCAGATGCTAAAAAT
GCCTTAATTGCTGGCGGCGTTGACGCTACCGATGCTAATGGCGCTGAGTTGGTCAAAATG
TCTTATACCGATAAAAATGGTAAGACAATTGAAGGCGGTTATGCGCTTAAAGCTGGCGAT
AAGTATTACGCCGCAGATTACGATGAAGCGACAGGAGCAATTAAAGCTAAAACTACAAGT

TATACTGCTGCTGACGGCACTACCAAAACAGCGGCTAACCAACTGGGTGGCGTAGACGGT
AAAACCGAAGTCGTTACTATCGACGGTAAAACCTACAATGCCAGCAAAGCCGCTGGTCAT
GATTTCAAAGCACAACCAGAGCTGGCGGAAGCAGCCGCTAAAACCACCGAAAACCCGCTG
CAGAAAATTGATGCCGCGCTGGCGCAGGTGGATGCGCTGCGCTCTGATCTGGGTGCGGTA
CAAAACCGTTTCAACTCTGCTATCACCAACCTGGGCAATACCGTAAACAATCTGTCTGAA
GCGCGTAGCCGTATCGAAGATTCCGACTACGCGACCGAAGTTTCCAACATGTCTCGCGCG
CAGATTCTGCAGCAGGCCGGTACTTCCGTTCTGGCGCAGGCTAACCAGGTCCCGCAGAAC
GTGCTGTCTCTGTTACGTAGCGGCAGTGGTAGCGGATCC

SEQ ID NO: 53 STF2.HA1-1 PR8

ATGGCACAAGTAATCAACACTAACAGTCTGTCGCTGCTGACCCAGAATAACCTGAACAAA
TCCCAGTCCGCACTGGGCACCGCTATCGAGCGTCTGTCTTCTGGTCTGCGTATCAACAGC
GCGAAAGACGATGCGGCAGGTCAGGCGATTGCTAACCGTTTCACCGCGAACATCAAAGGT
CTGACTCAGGCTTCCCGTAACGCTAACGACGGTATCTCCATTGCGCAGACCACTGAAGGC
GCGCTGAACGAAATCAACAACAACCTGCAGCGTGTGCGTGAACTGGCGGTTCAGTCTGCT
AACAGCACCAACTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATCACCCAGCGCCTG
AACGAAATCGACCGTGTATCCGGCCAGACTCAGTTCAACGGCGTGAAAGTCCTGGCGCAG
GACAACACCCTGACCATCCAGGTTGGCGCCAACGACGGTGAAACTATCGATATCGATCTG
AAGCAGATCAACTCTCAGACCCTGGGTCTGGACTCACTGAACGTGCAGAAAGCGTATGAT
GTGAAAGATACAGCAGTAACAACGAAAGCTTATGCCAATAATGGTACTACACTGGATGTA
TCGGGTCTTGATGATGCAGCTATTAAAGCGGCTACGGGTGGTACGAATGGTACGGCTTCT
GTAACCGGTGGTGCGGTTAAATTTGACGCAGATAATAACAAGTACTTTGTTACTATTGGT
GGCTTTACTGGTGCTGATGCCGCCAAAAATGGCGATTATGAAGTTAACGTTGCTACTGAC
GGTACAGTAACCCTTGCGGCTGGCGCAACTAAAACCACAATGCCTGCTGGTGCGACAACT
AAAACAGAAGTACAGGAGTTAAAAGATACACCGGCAGTTGTTTCAGCAGATGCTAAAAAT
GCCTTAATTGCTGGCGGCGTTGACGCTACCGATGCTAATGGCGCTGAGTTGGTCAAAATG
TCTTATACCGATAAAAATGGTAAGACAATTGAAGGCGGTTATGCGCTTAAAGCTGGCGAT
AAGTATTACGCCGCAGATTACGATGAAGCGACAGGAGCAATTAAAGCTAAAACCACAAGT
TATACTGCTGCTGACGGCACTACCAAAACAGCGGCTAACCAACTGGGTGGCGTAGACGGT
AAAACCGAAGTCGTTACTATCGACGGTAAAACCTACAATGCCAGCAAAGCCGCTGGTCAT
GATTTCAAAGCACAACCAGAGCTGGCGGAAGCAGCCGCTAAAACCACCGAAAACCCGCTG
CAGAAAATTGATGCCGCGCTGGCGCAGGTGGATGCGCTGCGCTCTGATCTGGGTGCGGTA
CAAAACCGTTTCAACTCTGCTATCACCAACCTGGGCAATACCGTAAACAATCTGTCTGAA
GCGCGTAGCCGTATCGAAGATTCCGACTACGCGACCGAAGTTTCCAACATGTCTCGCGCG
CAGATTTTGCAGCAGGCCGGTACTTCCGTTCTGGCGCAGGCTAACCAGGTCCCGCAGAAC
GTGCTGTCTCTGTTACGTAGCGGCAGTGGTAGCGGATCCTCTCATAACGGTAAACTGTGT
CGTCTGAAAGGTATTGCACCACTGCAGCTGGGCAAATGCAACATCGCGGGTTGGCTGCTG
GGTAACCCTGAATGTGACCCGCTGCTGCCGGTTCGTTCCTGGAGCTACATTGTTGAAACC
CCGAACTCCGAAAACGGTATCTGCTACCCGGGCGACTTTATTGACTATGAAGAACTGCGT
GAGCAGCTGTCTTCCGTGAGCAGCTTTGAACGCTTCGAAATCTTCCCGAAGGAAAGCTCC
TGGCCGAACCACAACACTAACGGCGTGACGGCGGCTTGCTCCCACGAAGGCAAATCTTCC
TTTTATCGTAACCTGCTGTGGCTGACTGAAAAGGAAGGTTCCTACCCAAAACTGAAAAAC
AGCTATGTTAACAAAAAGGGTAAAGAAGTCCTGGTGCTGTGGGGCATCCACCACCCGCCG
AACTCCAAGGAACAGCAGAATCTGTATCAGAACGAAAACGCATACGTTTCTGTCGTTACT
TCCAACTATAACCGTCGTTTCACTCCGGAAATCGCGGAACGTCCGAAAGTACGCGACCAG
GCTGGCCGTATGAACTACTACTGGACCCTGCTGAAACCGGGTGACACTATTATCTTCGAA
GCTAACGGTAACCTGATCGCACCAATGTACGCTTTCGCACTGTCTCGTGGTTTCGGTTCC
GGCATTATCACCAGCAACGCGTCTATGCACGAATGCAACACGAAATGTCAGACGCCGCTG
GGCGCAATTAATAGCAGCCTGCCGTACCAGAACATCCACCCGGTGACTATCGGCGAATGC
CCGAAGTATGTTCGTTAATAG

SEQ ID NO: 54 STF2.HA1-1.his PR8

ATGGCACAAGTAATCAACACTAACAGTCTGTCGCTGCTGACCCAGAATAACCTGAACAAA
TCCCAGTCCGCACTGGGCACCGCTATCGAGCGTCTGTCTTCTGGTCTGCGTATCAACAGC
GCGAAAGACGATGCGGCAGGTCAGGCGATTGCTAACCGTTTCACCGCGAACATCAAAGGT
CTGACTCAGGCTTCCCGTAACGCTAACGACGGTATCTCCATTGCGCAGACCACTGAAGGC
GCGCTGAACGAAATCAACAACAACCTGCAGCGTGTGCGTGAACTGGCGGTTCAGTCTGCT
AACAGCACCAACTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATCACCCAGCGCCTG
AACGAAATCGACCGTGTATCCGGCCAGACTCAGTTCAACGGCGTGAAAGTCCTGGCGCAG
GACAACACCCTGACCATCCAGGTTGGCGCCAACGACGGTGAAACTATCGATATCGATCTG
AAGCAGATCAACTCTCAGACCCTGGGTCTGGACTCACTGAACGTGCAGAAAGCGTATGAT
GTGAAAGATACAGCAGTAACAACGAAAGCTTATGCCAATAATGGTACTACACTGGATGTA
TCGGGTCTTGATGATGCAGCTATTAAAGCGGCTACGGGTGGTACGAATGGTACGGCTTCT
GTAACCGGTGGTGCGGTTAAATTTGACGCAGATAATAACAAGTACTTTGTTACTATTGGT
GGCTTTACTGGTGCTGATGCCGCCAAAAATGGCGATTATGAAGTTAACGTTGCTACTGAC
GGTACAGTAACCCTTGCGGCTGGCGCAACTAAAACCACAATGCCTGCTGGTGCGACAACT
AAAACAGAAGTACAGGAGTTAAAAGATACACCGGCAGTTGTTTCAGCAGATGCTAAAAAT
GCCTTAATTGCTGGCGGCGTTGACGCTACCGATGCTAATGGCGCTGAGTTGGTCAAAATG
TCTTATACCGATAAAAATGGTAAGACAATTGAAGGCGGTTATGCGCTTAAAGCTGGCGAT
AAGTATTACGCCGCAGATTACGATGAAGCGACAGGAGCAATTAAAGCTAAAACCACAAGT
TATACTGCTGCTGACGGCACTACCAAAACAGCGGCTAACCAACTGGGTGGCGTAGACGGT
AAAACCGAAGTCGTTACTATCGACGGTAAAACCTACAATGCCAGCAAAGCCGCTGGTCAT
GATTTCAAAGCACAACCAGAGCTGGCGGAAGCAGCCGCTAAAACCACCGAAAACCCGCTG
CAAAAAATTGATGCCGCGCTGGCGCAGGTGGATGCGCTGCGCTCTGATCTGGGTGCGGTA
CAAAACCGTTTCAACTCTGCTATCACCAACCTGGGCAATACCGTAAACAATCTGTCTGAA
GCGCGTAGCCGTATCGAAGATTCCGACTACGCGACCGAAGTTTCCAACATGTCTCGCGCG
CAGATTTTGCAGCAGGCCGGTACTTCCGTTCTGGCGCAGGCTAACCAGGTCCCGCAGAAC
GTGCTGTCTCTGTTACGTAGCGGCAGTGGTAGCGGATCCTCTCATAACGGTAAACTGTGT
CGTCTGAAAGGTATTGCACCACTGCAGCTGGGCAAATGCAACATCGCGGGTTGGCTGCTG
GGTAACCCTGAATGTGACCCGCTGCTGCCGGTTCGTTCCTGGAGCTACATTGTTGAAACC
CCGAACTCCGAAAACGGTATCTGCTACCCGGGCGACTTTATTGACTATGAAGAACTGCGT
GAGCAGCTGTCTTCCGTGAGCAGCTTTGAACGCTTCGAAATCTTCCCGAAGGAAAGCTCC
TGGCCGAACCACAACACTAACGGCGTGACGGCGGCTTGCTCCCACGAAGGCAAATCTTCC
TTTTATCGTAACCTGCTGTGGCTGACTGAAAAGGAAGGTTCCTACCCAAAACTGAAAAAC
AGCTATGTTAACAAAAAGGGTAAAGAAGTCCTGGTGCTGTGGGGCATCCACCACCCGCCG
AACTCCAAGGAACAGCAGAATCTGTATCAGAACGAAAACGCATACGTTTCTGTCGTTACT
TCCAACTATAACCGTCGTTTCACTCCGGAAATCGCGGAACGTCCGAAAGTACGCGACCAG
GCTGGCCGTATGAACTACTACTGGACCCTGCTGAAACCGGGTGACACTATTATCTTCGAA
GCTAACGGTAACCTGATCGCACCAATGTACGCTTTCGCACTGTCTCGTGGTTTCGGTTCC
GGCATTATCACCAGCAACGCGTCTATGCACGAATGCAACACGAAATGTCAGACGCCGCTG
GGCGCAATTAATAGCAGCCTGCCGTACCAGAACATCCACCCGGTGACTATCGGCGAATGC
CCGAAGTATGTTCGTCATCACCATCATCACCATTAATAG

SEQ ID NO: 55 STF2.HA1-2 PR8

ATGGCACAAGTAATCAACACTAACAGTCTGTCGCTGCTGACCCAGAATAACCTGAACAAA
TCCCAGTCCGCACTGGGCACCGCTATCGAGCGTCTGTCTTCTGGTCTGCGTATCAACAGC
GCGAAAGACGATGCGGCAGGTCAGGCGATTGCTAACCGTTTCACCGCGAACATCAAAGGT
CTGACTCAGGCTTCCCGTAACGCTAACGACGGTATCTCCATTGCGCAGACCACTGAAGGC
GCGCTGAACGAAATCAACAACAACCTGCAGCGTGTGCGTGAACTGGCGGTTCAGTCTGCT
AACAGCACCAACTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATCACCCAGCGCCTG
AACGAAATCGACCGTGTATCCGGCCAGACTCAGTTCAACGGCGTGAAAGTCCTGGCGCAG
GACAACACCCTGACCATCCAGGTTGGCGCCAACGACGGTGAAACTATCGATATCGATCTG
AAGCAGATCAACTCTCAGACCCTGGGTCTGGACTCACTGAACGTGCAGAAAGCGTATGAT
GTGAAAGATACAGCAGTAACAACGAAAGCTTATGCCAATAATGGTACTACACTGGATGTA

TCGGGTCTTGATGATGCAGCTATTAAAGCGGCTACGGGTGGTACGAATGGTACGGCTTCT
GTAACCGGTGGTGCGGTTAAATTTGACGCAGATAATAACAAGTACTTTGTTACTATTGGT
GGCTTTACTGGTGCTGATGCCGCCAAAAATGGCGATTATGAAGTTAACGTTGCTACTGAC
GGTACAGTAACCCTTGCGGCTGGCGCAACTAAAACCACAATGCCTGCTGGTGCGACAACT
AAAACAGAAGTACAGGAGTTAAAAGATACACCGGCAGTTGTTTCAGCAGATGCTAAAAAT
GCCTTAATTGCTGGCGGCGTTGACGCTACCGATGCTAATGGCGCTGAGTTGGTCAAAATG
TCTTATACCGATAAAAATGGTAAGACAATTGAAGGCGGTTATGCGCTTAAAGCTGGCGAT
AAGTATTACGCCGCAGATTACGATGAAGCGACAGGAGCAATTAAAGCTAAAACTACAAGT
TATACTGCTGCTGACGGCACTACCAAAACAGCGGCTAACCAACTGGGTGGCGTAGACGGT
AAAACCGAAGTCGTTACTATCGACGGTAAAACCTACAATGCCAGCAAAGCCGCTGGTCAT
GATTTCAAAGCACAACCAGAGCTGGCGGAAGCAGCCGCTAAAACCACCGAAAACCCGCTG
CAGAAAATTGATGCCGCGCTGGCGCAGGTGGATGCGCTGCGCTCTGATCTGGGTGCGGTA
CAAAACCGTTTCAACTCTGCTATCACCAACCTGGGCAATACCGTAAACAATCTGTCTGAA
GCGCGTAGCCGTATCGAAGATTCCGACTACGCGACCGAAGTTTCCAACATGTCTCGCGCG
CAGATTCTGCAGCAGGCCGGTACTTCCGTTCTGGCGCAGGCTAACCAGGTCCCGCAGAAC
GTGCTGTCTCTGTTACGTTCTGGTTCTGGTTCTGGTTCTAAAGGTATTGCTCCACTGCAA
CTGGGTAAATGCAATATTGCGGGTTGGCTGCTGGGCAACCCGGAATGCGATCCGCTGCTG
CCGGTCCGTTCCTGGAGCTATATTGTTGAAACTCCGAACTCTGAGAACGGCATCTGCTAT
CCAGGTGATTTCATTGACTATGAGGAACTGCGTGAACAACTGTCTTCCGTGTCTTCCTTT
GAACGTTTCGAGATTTTTCCTAAAGAATCTTCTTGGCCGAACCATAACACTAATGGTGTT
ACCGCTGCGTGCTCTCATGAAGGTAAATCTAGCTTTTACCGCAACCTGCTGTGGCTGACC
GAGAAAGAAGGTTCTTACCCGAAACTGAAAAACAGCTACGTAAACAAAAAGGGCAAGGAA
GTTCTGGTCCTGTGGGGTATCCACCATCCGCCGAACAGCAAGGAACAGCAGAATCTGTAT
CAGAACGAAAACGCATACGTATCTGTTGTTACTTCTAACTACAACCGCCGTTTCACCCCT
GAAATCGCGGAACGTCCGAAAGTGCGTGACCAGGCAGGCCGCATGAACTATTACTGGACC
CTGCTGAAGCCGGGTGATACTATCATCTTCGAAGCGAACGGTAACCTGATCGCCCCGATG
TACGCGTTCGCTCTGAGCCGTGGCTTCGGCTCTGGTATCATTACGTCTTAATAA

SEQ ID NO: 56 STF2.HA1-2*mut* PR8

ATGGCACAAGTAATCAACACTAACAGTCTGTCGCTGCTGACCCAGAATAACCTGAACAAA
TCCCAGTCCGCACTGGGCACCGCTATCGAGCGTCTGTCTTCTGGTCTGCGTATCAACAGC
GCGAAAGACGATGCGGCAGGTCAGGCGATTGCTAACCGTTTCACCGCGAACATCAAAGGT
CTGACTCAGGCTTCCCGTAACGCTAACGACGGTATCTCCATTGCGCAGACCACTGAAGGC
GCGCTGAACGAAATCAACAACAACCTGCAGCGTGTGCGTGAACTGGCGGTTCAGTCTGCT
AACAGCACCAACTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATCACCCAGCGCCTG
AACGAAATCGACCGTGTATCCGGCCAGACTCAGTTCAACGGCGTGAAAGTCCTGGCGCAG
GACAACACCCTGACCATCCAGGTTGGCGCCAACGACGGTGAAACTATCGATATCGATCTG
AAGCAGATCAACTCTCAGACCCTGGGTCTGGACTCACTGAACGTGCAGAAAGCGTATGAT
GTGAAAGATACAGCAGTAACAACGAAAGCTTATGCCAATAATGGTACTACACTGGATGTA
TCGGGTCTTGATGATGCAGCTATTAAAGCGGCTACGGGTGGTACGAATGGTACGGCTTCT
GTAACCGGTGGTGCGGTTAAATTTGACGCAGATAATAACAAGTACTTTGTTACTATTGGT
GGCTTTACTGGTGCTGATGCCGCCAAAAATGGCGATTATGAAGTTAACGTTGCTACTGAC
GGTACAGTAACCCTTGCGGCTGGCGCAACTAAAACCACAATGCCTGCTGGTGCGACAACT
AAAACAGAAGTACAGGAGTTAAAAGATACACCGGCAGTTGTTTCAGCAGATGCTAAAAAT
GCCTTAATTGCTGGCGGCGTTGACGCTACCGATGCTAATGGCGCTGAGTTGGTCAAAATG
TCTTATACCGATAAAAATGGTAAGACAATTGAAGGCGGTTATGCGCTTAAAGCTGGCGAT
AAGTATTACGCCGCAGATTACGATGAAGCGACAGGAGCAATTAAAGCTAAAACCACAAGT
TATACTGCTGCTGACGGCACTACCAAAACAGCGGCTAACCAACTGGGTGGCGTAGACGGT
AAAACCGAAGTCGTTACTATCGACGGTAAAACCTACAATGCCAGCAAAGCCGCTGGTCAT
GATTTCAAAGCACAACCAGAGCTGGCGGAAGCAGCCGCTAAAACCACCGAAAACCCGCTG
CAGAAAATTGATGCCGCGCTGGCGCAGGTGGATGCGCTGCGCTCTGATCTGGGTGCGGTA
CAAAACCGTTTCAACTCTGCTATCACCAACCTGGGCAATACCGTAAACAATCTGTCTGAA
GCGCGTAGCCGTATCGAAGATTCCGACTACGCGACCGAAGTTTCCAACATGTCTCGCGCG

CAGATTTTGCAGCAGGCCGGTACTTCCGTTCTGGCGCAGGCTAACCAGGTCCCGCAGAAC
GTGCTGTCTCTGTTAGCGAGCGGCAGTGGTAGCGGATCCAAAGGTGCGGCTCCACTGCAA
CTGGGTAAATGCAATATTGCGGGTTGGCTGCTGGGCAACCCGGAATGCGATCCGCTGCTG
CCGGTCCGTTCCTGGAGCGATATTGCGGAAACTCCGAACTCTGAGAACGGCATCTGCTAT
CCAGGTGATTTCATTGACTATGAGGAACTGCGTGAACAACTGTCTTCCGTGTCTTCCTTT
GAACGTTTCGAGATTTTTCCTAAAGAATCTTCTTGGCCGAACCATAACACTAATGGTGTT
ACCGCTGCGTGCTCTCATGAAGGTAAATCTAGCTTTTACCGCAACCTGCTGTGGCTGACC
GAGAAAGAAGGTTCTTACCCGAAACTGAAAAACAGCTACGTAAACAAAAAGGGCAAGGAA
GTTCTGGTCCTGTGGGGTATCCACCATCCGCCGAACAGCAAGGAACAGCAGAATCTGTAT
CAGAACGAAAACGCATACGTATCTGTTGTTACTTCTAACTACAACCGCCGTTTCACCCCT
GAAATCGCGGAACGTCCGAAAGTGCGTGACCAGGCAGGCCGCATGAACTATTACTGGACC
CTGCTGAAGCCGGGTGATACTATCATCTTCGAAGCGAACGGTAACCTGATCGCCCCGATG
TACGCGTTCGCTCTGAGCCGTGGCTTCGGCTCTGGTATCATTACGTCTTAATAA

SEQ ID NO: 57 STF2.HA1-3 PR8

ATGGCACAAGTAATCAACACTAACAGTCTGTCGCTGCTGACCCAGAATAACCTGAACAAA
TCCCAGTCCGCACTGGGCACCGCTATCGAGCGTCTGTCTTCTGGTCTGCGTATCAACAGC
GCGAAAGACGATGCGGCAGGTCAGGCGATTGCTAACCGTTTCACCGCGAACATCAAAGGT
CTGACTCAGGCTTCCCGTAACGCTAACGACGGTATCTCCATTGCGCAGACCACTGAAGGC
GCGCTGAACGAAATCAACAACAACCTGCAGCGTGTGCGTGAACTGGCGGTTCAGTCTGCT
AACAGCACCAACTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATCACCCAGCGCCTG
AACGAAATCGACCGTGTATCCGGCCAGACTCAGTTCAACGGCGTGAAAGTCCTGGCGCAG
GACAACACCCTGACCATCCAGGTTGGCGCCAACGACGGTGAAACTATCGATATCGATCTG
AAGCAGATCAACTCTCAGACCCTGGGTCTGGACTCACTGAACGTGCAGAAAGCGTATGAT
GTGAAAGATACAGCAGTAACAACGAAAGCTTATGCCAATAATGGTACTACACTGGATGTA
TCGGGTCTTGATGATGCAGCTATTAAAGCGGCTACGGGTGGTACGAATGGTACGGCTTCT
GTAACCGGTGGTGCGGTTAAATTTGACGCAGATAATAACAAGTACTTTGTTACTATTGGT
GGCTTTACTGGTGCTGATGCCGCCAAAAATGGCGATTATGAAGTTAACGTTGCTACTGAC
GGTACAGTAACCCTTGCGGCTGGCGCAACTAAAACCACAATGCCTGCTGGTGCGACAACT
AAAACAGAAGTACAGGAGTTAAAAGATACACCGGCAGTTGTTTCAGCAGATGCTAAAAAT
GCCTTAATTGCTGGCGGCGTTGACGCTACCGATGCTAATGGCGCTGAGTTGGTCAAAATG
TCTTATACCGATAAAAATGGTAAGACAATTGAAGGCGGTTATGCGCTTAAAGCTGGCGAT
AAGTATTACGCCGCAGATTACGATGAAGCGACAGGAGCAATTAAAGCTAAAACTACAAGT
TATACTGCTGCTGACGGCACTACCAAAACAGCGGCTAACCAACTGGGTGGCGTAGACGGT
AAAACCGAAGTCGTTACTATCGACGGTAAAACCTACAATGCCAGCAAAGCCGCTGGTCAT
GATTTCAAAGCACAACCAGAGCTGGCGGAAGCAGCCGCTAAAACCACCGAAAACCCGCTG
CAGAAAATTGATGCCGCGCTGGCGCAGGTGGATGCGCTGCGCTCTGATCTGGGTGCGGTA
CAAAACCGTTTCAACTCTGCTATCACCAACCTGGGCAATACCGTAAACAATCTGTCTGAA
GCGCGTAGCCGTATCGAAGATTCCGACTACGCGACCGAAGTTTCCAACATGTCTCGCGCG
CAGATTCTGCAGCAGGCCGGTACTTCCGTTCTGGCGCAGGCTAACCAGGTCCCGCAGAAC
GTGCTGTCTCTGTTACGTTCTGGTAGCGGTTCTGGCTCTAACTCTGAAAATGGTATCTGT
TACCCGGGTGATTTCATCGATTATGAGGAACTGCGTGAACAGCTGTCTAGCGTGAGCTCC
TTTGAACGCTTCGAAATCTTCCCGAAGGAATCTAGCTGGCCGAACCATAACACGAACGGT
GTGACCGCTGCTTGCTCCCACGAAGGTAAGAGCTCCTTCTACCGCAATCTGCTGTGGCTG
ACTGAAAAAGAAGGTAGCTACCCGAAACTGAAAAATTCTTACGTCAACAAGAAAGGCAAG
GAAGTGCTGGTTCTGTGGGGCATTCACCACCCACCGAACAGCAAAGAGCAACAGAACCTG
TACCAAAATGAGAACGCTTACGTTTCTGTTGTGACTTCTAACTACAATCGTCGCTTTACC
CCTGAAATCGCGGAGCGTCCAAAAGTGCGTGACCAGGCTGGTCGTATGAACTACTATTGG
ACCCTGCTGAAACCGGGCGACACCATTATCTTTGAAGCGAACGGTAACCTGATCGCGCCT
ATGTACGCGTTCGCTCTGTCTCGTGGCTAATAA

SEQ ID NO: 58 STF2.HA1-3mut PR8

ATGGCACAAGTAATCAACACTAACAGTCTGTCGCTGCTGACCCAGAATAACCTGAACAAA
TCCCAGTCCGCACTGGGCACCGCTATCGAGCGTCTGTCTTCTGGTCTGCGTATCAACAGC
GCGAAAGACGATGCGGCAGGTCAGGCGATTGCTAACCGTTTCACCGCGAACATCAAAGGT
CTGACTCAGGCTTCCCGTAACGCTAACGACGGTATCTCCATTGCGCAGACCACTGAAGGC
GCGCTGAACGAAATCAACAACAACCTGCAGCGTGTGCGTGAACTGGCGGTTCAGTCTGCT
AACAGCACCAACTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATCACCCAGCGCCTG
AACGAAATCGACCGTGTATCCGGCCAGACTCAGTTCAACGGCGTGAAAGTCCTGGCGCAG
GACAACACCCTGACCATCCAGGTTGGCGCCAACGACGGTGAAACTATCGATATCGATCTG
AAGCAGATCAACTCTCAGACCCTGGGTCTGGACTCACTGAACGTGCAGAAAGCGTATGAT
GTGAAAGATACAGCAGTAACAACGAAAGCTTATGCCAATAATGGTACTACACTGGATGTA
TCGGGTCTTGATGATGCAGCTATTAAAGCGGCTACGGGTGGTACGAATGGTACGGCTTCT
GTAACCGGTGGTGCGGTTAAATTTGACGCAGATAATAACAAGTACTTTGTTACTATTGGT
GGCTTTACTGGTGCTGATGCCGCCAAAAATGGCGATTATGAAGTTAACGTTGCTACTGAC
GGTACAGTAACCCTTGCGGCTGGCGCAACTAAAACCACAATGCCTGCTGGTGCGACAACT
AAAACAGAAGTACAGGAGTTAAAAGATACACCGGCAGTTGTTTCAGCAGATGCTAAAAAT
GCCTTAATTGCTGGCGGCGTTGACGCTACCGATGCTAATGGCGCTGAGTTGGTCAAAATG
TCTTATACCGATAAAAATGGTAAGACAATTGAAGGCGGTTATGCGCTTAAAGCTGGCGAT
AAGTATTACGCCGCAGATTACGATGAAGCGACAGGAGCAATTAAAGCTAAAACTACAAGT
TATACTGCTGCTGACGGCACTACCAAAACAGCGGCTAACCAACTGGGTGGCGTAGACGGT
AAAACCGAAGTCGTTACTATCGACGGTAAAACCTACAATGCCAGCAAAGCCGCTGGTCAT
GATTTCAAAGCACAACCAGAGCTGGCGGAAGCAGCCGCTAAAACCACCGAAAACCCGCTG
CAGAAAATTGATGCCGCGCTGGCGCAGGTGGATGCGCTGCGCTCTGATCTGGGTGCGGTA
CAAAACCGTTTCAACTCTGCTATCACCAACCTGGGCAATACCGTAAACAATCTGTCTGAA
GCGCGTAGCCGTATCGAAGATTCCGACTACGCGACCGAAGTTTCCAACATGTCTCGCGCG
CAGATTCTGCAGCAGGCCGGTACTTCCGTTCTGGCGCAGGCTAACCAGGTCCCGCAGAAC
GTGCTGTCTCTGTTACGTAGCGGCAGTGGTAGCGGATCCAACTCTGAAAATGAAATCTGT
TACCCGGGTGATTTCATCGATAAAGAGGAACTGCGTGAACAGCTGTCTAGCGTGAGCTCC
TTTGAACGCTTCGAAATCTTCCCGAAGGAATCTAGCTGGCCGAACCATAACACGAACGGT
GTGACCGCTGCTTGCTCCCACGAAGGTAAGAGCTCCTTCTACCGCAATCTGCTGTGGCTG
ACTGAAAAAGAAGGTAGCTACCCGAAACTGAAAAATTCTTACGTCAACAAGAAAGGCAAG
GAAGTGCTGGTTCTGTGGGGCATTCACCACCCACCGAACAGCAAAGAGCAACAGAACCTG
TACCAAAATGAGAACGCTTACGTTTCTGTTGTGACTTCTAACTACAATCGTCGCTTTACC
CCTGAAATCGCGGAGCGTCCAAAAGTGCGTGACCAGGCTGGTCGTATGAACTACTATTGG
ACCCTGCTGAAACCGGGCGACACCATTATCTTTGAAGCGAACGGTAACCTGATCGCGCCT
ATGTACGCGGCGGCTCTGTCTCGTGGCTAATAA

SEQ ID NO: 59 HA1-1 PR8

ATGTCTCATAACGGTAAACTGTGTCGTCTGAAAGGTATTGCACCACTGCAGCTGGGCAAA
TGCAACATCGCGGGTTGGCTGCTGGGTAACCCTGAATGTGACCCGCTGCTGCCGGTTCGT
TCCTGGAGCTACATTGTTGAAACCCCGAACTCCGAAAACGGTATCTGCTACCCGGGCGAC
TTTATTGACTATGAAGAACTGCGTGAGCAGCTGTCTTCCGTGAGCAGCTTTGAACGCTTC
GAAATCTTCCCGAAGGAAAGCTCCTGGCCGAACCACAACACTAACGGCGTGACGGCGGCT
TGCTCCCACGAAGGCAAATCTTCCTTTTATCGTAACCTGCTGTGGCTGACTGAAAAGGAA
GGTTCCTACCCAAAACTGAAAAACAGCTATGTTAACAAAAAGGGTAAAGAAGTCCTGGTG
CTGTGGGGCATCCACCACCCGCCGAACTCCAAGGAACAGCAGAATCTGTATCAGAACGAA
AACGCATACGTTTCTGTCGTTACTTCCAACTATAACCGTCGTTTCACTCCGGAAATCGCG
GAACGTCCGAAAGTACGCGACCAGGCTGGCCGTATGAACTACTACTGGACCCTGCTGAAA
CCGGGTGACACTATTATCTTCGAAGCTAACGGTAACCTGATCGCACCAATGTACGCTTTC
GCACTGTCTCGTGGTTTCGGTTCCGGCATTATCACCAGCAACGCGTCTATGCACGAATGC
AACACGAAATGTCAGACGCCGCTGGGCGCAATTAATAGCAGCCTGCCGTACCAGAACATC
CACCCGGTGACTATCGGCGAATGCCCGAAGTATGTTCGT

SEQ ID NO: 60 HA1-1.his PR8

ATGTCTCATAACGGTAAACTGTGTCGTCTGAAAGGTATTGCACCACTGCAGCTGGGCAAA
TGCAACATCGCGGGTTGGCTGCTGGGTAACCCTGAATGTGACCCGCTGCTGCCGGTTCGT
TCCTGGAGCTACATTGTTGAAACCCCGAACTCCGAAAACGGTATCTGCTACCCGGGCGAC
TTTATTGACTATGAAGAACTGCGTGAGCAGCTGTCTTCCGTGAGCAGCTTTGAACGCTTC
GAAATCTTCCCGAAGGAAAGCTCCTGGCCGAACCACAACACTAACGGCGTGACGGCGGCT
TGCTCCCACGAAGGCAAATCTTCCTTTTATCGTAACCTGCTGTGGCTGACTGAAAAGGAA
GGTTCCTACCCAAAACTGAAAAACAGCTATGTTAACAAAAAGGGTAAAGAAGTCCTGGTG
CTGTGGGGCATCCACCACCCGCCGAACTCCAAGGAACAGCAGAATCTGTATCAGAACGAA
AACGCATACGTTTCTGTCGTTACTTCCAACTATAACCGTCGTTTCACTCCGGAAATCGCG
GAACGTCCGAAAGTACGCGACCAGGCTGGCCGTATGAACTACTACTGGACCCTGCTGAAA
CCGGGTGACACTATTATCTTCGAAGCTAACGGTAACCTGATCGCACCAATGTACGCTTTC
GCACTGTCTCGTGGTTTCGGTTCCGGCATTATCACCAGCAACGCGTCTATGCACGAATGC
AACACGAAATGTCAGACGCCGCTGGGCGCAATTAATAGCAGCCTGCCGTACCAGAACATC
CACCCGGTGACTATCGGCGAATGCCCGAAGTATGTTCGTCATCACCATCATCACCAT

SEQ ID NO: 61 HA1-2.his PR8

ATGAAAGGTATTGCTCCACTGCAACTGGGTAAATGCAATATTGCGGGTTGGCTGCTGGGC
AACCCGGAATGCGATCCGCTGCTGCCGGTCCGTTCCTGGAGCTATATTGTTGAAACTCCG
AACTCTGAGAACGGCATCTGCTATCCAGGTGATTTCATTGACTATGAGGAACTGCGTGAA
CAACTGTCTTCCGTGTCTTCCTTTGAACGTTTCGAGATTTTTCCTAAAGAATCTTCTTGG
CCGAACCATAACACTAATGGTGTTACCGCTGCGTGCTCTCATGAAGGTAAATCTAGCTTT
TACCGCAACCTGCTGTGGCTGACCGAGAAAGAAGGTTCTTACCCGAAACTGAAAAACAGC
TACGTAAACAAAAAGGGCAAGGAAGTTCTGGTCCTGTGGGGTATCCACCATCCGCCGAAC
AGCAAGGAACAGCAGAATCTGTATCAGAACGAAAACGCATACGTATCTGTTGTTACTTCT
AACTACAACCGCCGTTTCACCCCTGAAATCGCGGAACGTCCGAAAGTGCGTGACCAGGCA
GGCCGCATGAACTATTACTGGACCCTGCTGAAGCCGGGTGATACTATCATCTTCGAAGCG
AACGGTAACCTGATCGCCCCGATGTACGCGTTCGCTCTGAGCCGTGGCTTCGGCTCTGGT
ATCATTACGTCTCATCACCATCATCACCAT

SEQ ID NO: 62 CRM.HA1-2

ATGGGTGCTGATGATGTCGTTGATTCCTCCAAAAGCTTCGTTATGGAAAATTTCTCTTCT
TATCACGGCACCAAACCGGGTTATGTGGATTCTATCCAAAAAGGCATCCAGAAACCGAAG
TCCGGTACGCAGGGTAACTATGATGACGATTGGAAAGGTTTTTACTCCACCGATAATAAA
TATGACGCGGCTGGCTACTCTGTTGACAACGAAAATCCACTGTCTGGTAAGGCTGGCGGT
GTCGTAAAGGTAACGTATCCTGGCCTGACCAAGGTGCTGGCACTGAAGGTGGATAACGCT
GAAACCATCAAAAAGGAGCTGGGCCTGAGCCTGACTGAACCGCTGATGGAGCAGGTCGGT
ACCGAGGAATTCATCAAACGCTTCGGTGATGGCGCCTCCCGTGTTGTGCTGTCCCTGCCG
TTCGCAGAAGGCTCTTCCTCTGTCGAATATATCAACAACTGGGAACAGGCTAAAGCTCTG
AGCGTCGAACTGGAAATTAACTTTGAGACCCGTGGCAAGCGTGGTCAGGACGCGATGTAC
GAATATATGGCTCAGGCTTGTGCCGGTAACCGTGTTCGCCGTTCCGTCGGTTCCTCTCTG
TCTTGCATCAACCTGGATTGGGACGTTATCCGTGATAAGACCAAAACCAAAATTGAAAGC
CTGAAGGAACACGGTCCGATCAAAAACAAAATGTCTGAATCTCCGAACAAAACCGTGTCC
GAGGAAAAAGCGAAACAGTATCTGGAAGAATTCCACCAGACTGCCCTGGAACATCCTGAA
CTGTCCGAACTGAAGACTGTAACCGGCACTAACCCGGTGTTCGCAGGCGCAAACTACGCC
GCGTGGGCGGTAAACGTTGCGCAGGTTATTGATAGCGAAACCGCAGATAACCTGGAAAAA
ACGACCGCAGCTCTGTCTATCCTGCCGGGTATCGGTTCCGTTATGGGTATTGCGGACGGC
GCTGTGCACCACAACACGGAAGAAATCGTCGCACAGTCTATCGCGCTGTCTTCTCTGATG
GTTGCTCAGGCAATTCCACTGGTAGGTGAACTGGTGGACATTGGCTTTGCGGCGTACAAC
TTCGTCGAAAGCATTATCAACCTGTTCCAGGTTGTACACAACTCTTACAACCGTCCGGCC
TACAGCCCTGGCCACAAAACCCAACCGTTTCTGCACGACGGTTATGCGGTGTCCTGGAAC
ACGGTTGAAGACTCTATCATTCGTACCGGCTTTCAGGGCGAGTCCGGCCACGATATCAAA
ATTACTGCAGAAAACACTCCGCTGCCGATCGCTGGCGTTCTGCTGCCGACCATCCCGGGT

AAGCTGGATGTGAACAAAAGCAAAACCCACATCTCTGTTAACGGTCGTAAAATTCGCATG
CGCTGTCGTGCTATCGACGGTGATGTTACCTTCTGCCGTCCGAAATCTCCAGTCTACGTG
GGCAACGGTGTTCATGCCAACCTGCACGTGGCGTTCCATCGTAGCTCTAGCGAAAAAATC
CACTCCAACGAAATCTCTAGCGATTCTATCGGTGTTCTGGGTTATCAGAAAACGGTGGAT
CATACGAAAGTCAATTCTAAACTGAGCCTGTTCTTCGAAATCAAATCTAGCGGCTCTGGA
TCCGGTTCCAAAGGCATCGCGCCGCTGCAGCTGGGTAAATGTAACATTGCGGGCTGGCTG
CTGGGTAATCCGGAATGCGATCCGCTGCTGCCGGTCCGTAGCTGGTCTTACATTGTTGAA
ACTCCGAACTCTGAGAATGGCATCTGCTACCCGGGCGATTTTATCGACTATGAAGAACTG
CGTGAACAGCTGTCTTCCGTTTCTTCCTTTGAACGTTTCGAAATCTTCCCGAAAGAAAGC
AGCTGGCCGAATCACAATACGAACGGTGTTACTGCTGCGTGTTCTCATGAAGGTAAATCC
AGCTTCTACCGTAACCTGCTGTGGCTGACCGAAAAAGAGGGTTCTTATCCTAAACTGAAA
AACAGCTACGTTAACAAAAAGGGCAAAGAAGTGCTGGTGCTGTGGGGTATCCATCACCCT
CCGAACTCTAAAGAACAACAGAATCTGTATCAGAACGAAAACGCTTACGTTTCCGTGGTG
ACCTCTAACTATAACCGTCGTTTTACCCCGGAGATTGCTGAACGTCCGAAAGTGCGCGAT
CAGGCTGGCCGTATGAACTACTATTGGACCCTGCTGAAACCGGGCGATACCATCATTTTC
GAAGCTAACGGCAACCTGATTGCTCCGATGTATGCGTTTGCTCTGTCTCGTGGCTTCGGC
TCTGGTATTATTACGTCTTAATAA

SEQ ID NO: 63 LTB.HA1-2

ATGAATAAGGTTAAGTTCTACGTACTGTTTACCGCGCTGCTGTCTTCTCTGTGCGCGCAT
GGTGCTCCGCAGTCTATTACTGAACTGTGCTCTGAATACCACAACACCCAGATCTATACT
ATCAACGATAAAATCCTGAGCTATACCGAATCTATGGCAGGCAAACGCGAAATGGTTATC
ATTACCTTTAAAAGCGGCGCCACCTTTCAAGTGGAAGTTCCGGGCTCTCAGCATATTGAC
TCTCAAAAAAAAGCGATCGAACGTATGAAAGATACTCTGCGCATTACCTACCTGACCGAA
ACCAAAATCGATAAACTGTGCGTATGGAACAATAAGACTCCTAACTCTATCGCAGCTATT
TCTATGGAAAACTCTGGTAGCGGATCCGGTTCTAAAGGCATCGCGCCGCTGCAGCTGGGT
AAATGTAACATTGCGGGCTGGCTGCTGGGTAATCCGGAATGCGATCCGCTGCTGCCGGTC
CGTAGCTGGTCTTACATTGTTGAAACTCCGAACTCTGAGAATGGCATCTGCTACCCGGGC
GATTTTATCGACTATGAAGAACTGCGTGAACAGCTGTCTTCCGTTTCTTCCTTTGAACGT
TTCGAAATCTTCCCGAAAGAAAGCAGCTGGCCGAATCACAATACGAACGGTGTTACTGCT
GCGTGTTCTCATGAAGGTAAATCCAGCTTCTACCGTAACCTGCTGTGGCTGACCGAAAAA
GAGGGTTCTTATCCTAAACTGAAAAACAGCTACGTTAACAAAAAGGGCAAAGAAGTGCTG
GTGCTGTGGGGTATCCATCACCCTCCGAACTCTAAAGAACAACAGAATCTGTATCAGAAC
GAAAACGCTTACGTTTCCGTGGTGACCTCTAACTATAACCGTCGTTTTACCCCGGAGATT
GCTGAACGTCCGAAAGTGCGCGATCAGGCTGGCCGTATGAACTACTATTGGACCCTGCTG
AAACCGGGCGATACCATCATTTTCGAAGCTAACGGCAACCTGATTGCTCCGATGTATGCG
TTTGCTCTGTCTCGTGGCTTCGGCTCTGGTATTATTACGTCTTAATAA

SEQ ID NO: 64 STF2.HA1-1 VN

ATGGCACAAGTAATCAACACTAACAGTCTGTCGCTGCTGACCCAGAATAACCTGAACAAA
TCCCAGTCCGCACTGGGCACCGCTATCGAGCGTCTGTCTTCTGGTCTGCGTATCAACAGC
GCGAAAGACGATGCGGCAGGTCAGGCGATTGCTAACCGTTTCACCGCGAACATCAAAGGT
CTGACTCAGGCTTCCCGTAACGCTAACGACGGTATCTCCATTGCGCAGACCACTGAAGGC
GCGCTGAACGAAATCAACAACAACCTGCAGCGTGTGCGTGAACTGGCGGTTCAGTCTGCT
AACAGCACCAACTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATCACCCAGCGCCTG
AACGAAATCGACCGTGTATCCGGCCAGACTCAGTTCAACGGCGTGAAAGTCCTGGCGCAG
GACAACACCCTGACCATCCAGGTTGGCGCCAACGACGGTGAAACTATCGATATCGATCTG
AAGCAGATCAACTCTCAGACCCTGGGTCTGGACTCACTGAACGTGCAGAAAGCGTATGAT
GTGAAAGATACAGCAGTAACAACGAAAGCTTATGCCAATAATGGTACTACACTGGATGTA
TCGGGTCTTGATGATGCAGCTATTAAAGCGGCTACGGGTGGTACGAATGGTACGGCTTCT
GTAACCGGTGGTGCGGTTAAATTTGACGCAGATAATAACAAGTACTTTGTTACTATTGGT
GGCTTTACTGGTGCTGATGCCGCCAAAAATGGCGATTATGAAGTTAACGTTGCTACTGAC
GGTACAGTAACCCTTGCGGCTGGCGCAACTAAAACCACAATGCCTGCTGGTGCGACAACT

AAAACAGAAGTACAGGAGTTAAAAGATACACCGGCAGTTGTTTCAGCAGATGCTAAAAAT
GCCTTAATTGCTGGCGGCGTTGACGCTACCGATGCTAATGGCGCTGAGTTGGTCAAAATG
TCTTATACCGATAAAAATGGTAAGACAATTGAAGGCGGTTATGCGCTTAAAGCTGGCGAT
AAGTATTACGCCGCAGATTACGATGAAGCGACAGGAGCAATTAAAGCTAAAACCACAAGT
TATACTGCTGCTGACGGCACTACCAAAACAGCGGCTAACCAACTGGGTGGCGTAGACGGT
AAAACCGAAGTCGTTACTATCGACGGTAAAACCTACAATGCCAGCAAAGCCGCTGGTCAT
GATTTCAAAGCACAACCAGAGCTGGCGGAAGCAGCCGCTAAAACCACCGAAAACCCGCTG
CAGAAAATTGATGCCGCGCTGGCGCAGGTGGATGCGCTGCGCTCTGATCTGGGTGCGGTA
CAAAACCGTTTCAACTCTGCTATCACCAACCTGGGCAATACCGTAAACAATCTGTCTGAA
GCGCGTAGCCGTATCGAAGATTCCGACTACGCGACCGAAGTTTCCAACATGTCTCGCGCG
CAGATTTTGCAGCAGGCCGGTACTTCCGTTCTGGCGCAGGCTAACCAGGTCCCGCAGAAC
GTGCTGAGCCTGTTAGCGGAGAAGAAACACAATGGCAAACTGTGTGATCTGGATGGTGTG
AAACCGCTGATTCTGCGCGATTGCTCTGTGGCAGGCTGGCTGCTGGGCAACCCTATGTGT
GACGAATTCATTAACGTTCCGGAATGGTCTTACATTGTTGAAAAAGCTAACCCTGTCAAC
GATCTGTGTTACCCTGGTGACTTTAACGATTACGAAGAACTGAAGCACCTGCTGTCTCGT
ATCAATCACTTCGAGAAAATCCAGATCATCCCGAAATCCTCCTGGAGCTCCCACGAAGCT
TCTCTGGGCGTATCCTCCGCGTGCCCGTACCAGGGCAAATCCTCTTTCTTTCGTAACGTT
GTTTGGCTGATCAAGAAAAACTCCACCTACCCGACGATCAAGCGTAGCTATAATAACACC
AACCAGGAAGACCTGCTGGTTCTGTGGGGCATCCACCATCCAAACGATGCTGCGGAACAG
ACCAAGCTGTACCAGAACCCGACCACCTACATCAGCGTGGGCACCTCTACGCTGAACCAG
CGTCTGGTACCGCGTATCGCAACCCGCAGCAAGGTAAACGGTCAAAGCGGCCGCATGGAA
TTTTTCTGGACCATCCTGAAACCGAACGACGCAATCAACTTCGAATCTAACGGCAATTTC
ATCGCTCCGGAGTATGCGTACAAAATCGTAAAGAAAGGTGATAGCACTATCATGAAATCC
GAGCTGGAATATGGCAACTGTAACACCAAATGCCAGACCCCGATGGGTGCAATCAACTCC
TCCATGCCGTTTCACAACATTCACCCGCTGACTATCGGCGAATGTCCGAAATACGTTAAA
TAGTAA

SEQ ID NO: 65 FOR PRIMER
AGTGGCTGAGCCTGTTAGCGGAGAAGAAACACAATGGCAAACTGTG

SEQ ID NO: 66 REV PRIMER
AGTCGCTCAGCTTACTATTTAACGTATTTCGGACATTCGCCGATAGTCAG

SEQ ID NO: 67 HA0s VN
GTGCTGAGCCTGTTACGTCAGATTTGTATCGGCTACCACGCAAACAACTCTACCGAGCAA
GTTGATACCATCATGGAGAAAAACGTGACCGTTACTCACGCGCAGGACATCCTGGAGAAG
AAACACAATGGCAAACTGTGTGATCTGGATGGTGTGAAACCGCTGATTCTGCGCGATTGC
TCTGTGGCAGGCTGGCTGCTGGGCAACCCTATGTGTGACGAATTCATTAACGTTCCGGAA
TGGTCTTACATTGTTGAAAAAGCTAACCCTGTCAACGATCTGTGTTACCCTGGTGACTTT
AACGATTACGAAGAACTGAAGCACCTGCTGTCTCGTATCAATCACTTCGAGAAAATCCAG
ATCATCCCGAAATCCTCCTGGAGCTCCCACGAAGCTTCTCTGGGCGTATCCTCCGCGTGC
CCGTACCAGGGCAAATCCTCTTTCTTTCGTAACGTTGTTTGGCTGATCAAGAAAAACTCC
ACCTACCCGACGATCAAGCGTAGCTATAATAACACCAACCAGGAAGACCTGCTGGTTCTG
TGGGGCATCCACCATCCAAACGATGCTGCGGAACAGACCAAGCTGTACCAGAACCCGACC
ACCTACATCAGCGTGGGCACCTCTACGCTGAACCAGCGTCTGGTACCGCGTATCGCAACC
CGCAGCAAGGTAAACGGTCAAAGCGGCCGCATGGAATTTTTCTGGACCATCCTGAAACCG
AACGACGCAATCAACTTCGAATCTAACGGCAATTTCATCGCTCCGGAGTATGCGTACAAA
ATCGTAAAGAAAGGTGATAGCACTATCATGAAATCCGAGCTGGAATATGGCAACTGTAAC
ACCAAATGCCAGACCCCGATGGGTGCAATCAACTCCTCCATGCCGTTTCACAACATTCAC
CCGCTGACTATCGGCGAATGTCCGAAATACGTTAAATCCAATCGTCTGGTTCTGGCTACC
GGTCTGCGTAACTCCCCACAGCGTGAACGTCGTCGTAAGAAACGTGGTCTGTTTGGCGCG
ATCGCTGGTTTCATCGAGGGCGGCTGGCAGGGTATGGTTGATGGCTGGTACGGTTATCAT
CATTCCAATGAACAGGGTTCCGGCTACGCCGCAGATAAAGAAAGCACTCAGAAAGCAATT

GATGGCGTAACTAACAAAGTAAATTCTATCATTGATAAAATGAACACCCAGTTCGAGGCG
GTTGGTCGTGAGTTCAACAACCTGGAACGCCGTATCGAAAACCTGAACAAGAAAATGGAA
GACGGTTTTCTGGATGTGTGGACTTACAATGCTGAACTGCTGGTGCTGATGGAAAACGAG
CGTACCCTGGACTTCCACGACAGCAACGTCAAAAATCTGTATGACAAAGTCCGTCTGCAG
CTGCGTGATAACGCTAAAGAGCTGGGTAATGGCTGCTTCGAGTTCTATCACAAATGCGAC
AACGAATGCATGGAATCTGTCCGCAACGGCACTTACGATTATCCGCAGTACTCCGAAGAA
GCGCGCCTGAAACGCGAAGAGATCTCCGGTGTGAAGCTGGAGTCTATTGGCATCTACCAG
ATCCTGTCCATCTACAGCACCTAGTAAGCTGAGCGCCTACGCAGC

SEQ ID NO: 68 STF2.HA1-2 VN

ATGGCACAAGTAATCAACACTAACAGTCTGTCGCTGCTGACCCAGAATAACCTGAACAAA
TCCCAGTCCGCACTGGGCACCGCTATCGAGCGTCTGTCTTCTGGTCTGCGTATCAACAGC
GCGAAAGACGATGCGGCAGGTCAGGCGATTGCTAACCGTTTCACCGCGAACATCAAAGGT
CTGACTCAGGCTTCCCGTAACGCTAACGACGGTATCTCCATTGCGCAGACCACTGAAGGC
GCGCTGAACGAAATCAACAACAACCTGCAGCGTGTGCGTGAACTGGCGGTTCAGTCTGCT
AACAGCACCAACTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATCACCCAGCGCCTG
AACGAAATCGACCGTGTATCCGGCCAGACTCAGTTCAACGGCGTGAAAGTCCTGGCGCAG
GACAACACCCTGACCATCCAGGTTGGCGCCAACGACGGTGAAACTATCGATATCGATCTG
AAGCAGATCAACTCTCAGACCCTGGGTCTGGACTCACTGAACGTGCAGAAAGCGTATGAT
GTGAAAGATACAGCAGTAACAACGAAAGCTTATGCCAATAATGGTACTACACTGGATGTA
TCGGGTCTTGATGATGCAGCTATTAAAGCGGCTACGGGTGGTACGAATGGTACGGCTTCT
GTAACCGGTGGTGCGGTTAAATTTGACGCAGATAATAACAAGTACTTTGTTACTATTGGT
GGCTTTACTGGTGCTGATGCCGCCAAAAATGGCGATTATGAAGTTAACGTTGCTACTGAC
GGTACAGTAACCCTTGCGGCTGGCGCAACTAAAACCACAATGCCTGCTGGTGCGACAACT
AAAACAGAAGTACAGGAGTTAAAAGATACACCGGCAGTTGTTTCAGCAGATGCTAAAAAT
GCCTTAATTGCTGGCGGCGTTGACGCTACCGATGCTAATGGCGCTGAGTTGGTCAAAATG
TCTTATACCGATAAAAATGGTAAGACAATTGAAGGCGGTTATGCGCTTAAAGCTGGCGAT
AAGTATTACGCCGCAGATTACGATGAAGCGACAGGAGCAATTAAAGCTAAAACCACAAGT
TATACTGCTGCTGACGGCACTACCAAAACAGCGGCTAACCAACTGGGTGGCGTAGACGGT
AAAACCGAAGTCGTTACTATCGACGGTAAAACCTACAATGCCAGCAAAGCCGCTGGTCAT
GATTTCAAAGCACAACCAGAGCTGGCGGAAGCAGCCGCTAAAACCACCGAAAACCCGCTG
CAGAAAATTGATGCCGCGCTGGCGCAGGTGGATGCGCTGCGCTCTGATCTGGGTGCGGTA
CAAAACCGTTTCAACTCTGCTATCACCAACCTGGGCAATACCGTAAACAATCTGTCTGAA
GCGCGTAGCCGTATCGAAGATTCCGACTACGCGACCGAAGTTTCCAACATGTCTCGCGCG
CAGATTTTGCAGCAGGCCGGTACTTCCGTTCTGGCGCAGGCTAACCAGGTCCCGCAGAAC
GTGCTGAGCCTGTTAGCAGGTGTGAAACCGCTGATTCTGCGCGATTGCTCTGTGGCAGGC
TGGCTGCTGGGCAACCCTATGTGTGACGAATTCATTAACGTTCCGGAATGGTCTTACATT
GTTGAAAAAGCTAACCCTGTCAACGATCTGTGTTACCCTGGTGACTTTAACGATTACGAA
GAACTGAAGCACCTGCTGTCTCGTATCAATCACTTCGAGAAAATCCAGATCATCCCGAAA
TCCTCCTGGAGCTCCCACGAAGCTTCTCTGGGCGTATCCTCCGCGTGCCCGTACCAGGGC
AAATCCTCTTTCTTTCGTAACGTTGTTTGGCTGATCAAGAAAAACTCCACCTACCCGACG
ATCAAGCGTAGCTATAATAACACCAACCAGGAAGACCTGCTGGTTCTGTGGGGCATCCAC
CATCCAAACGATGCTGCGGAACAGACCAAGCTGTACCAGAACCCGACCACCTACATCAGC
GTGGGCACCTCTACGCTGAACCAGCGTCTGGTACCGCGTATCGCAACCCGCAGCAAGGTA
AACGGTCAAAGCGGCCGCATGGAATTTTTCTGGACCATCCTGAAACCGAACGACGCAATC
AACTTCGAATCTAACGGCAATTTCATCGCTCCGGAGTATGCGTACAAAATCGTAAAGAAA
GGTGATAGCACTATCATGAAATCCGAGTAGTAA

SEQ ID NO: 69 FOR PRIMER

AGTCGCTGAGCCTGTTAGCAGGTGTGAAACCGCTGATTCTGCGCGATTG

SEQ ID NO: 70 REV PRIMER

TGACGCTCAGCTTACTACTCGGATTTCATGATAGTGCTATCACCTTTC

SEQ ID NO: 71 STF2.HA1-2mut VN

ATGGCACAAGTAATCAACACTAACAGTCTGTCGCTGCTGACCCAGAATAACCTGAACAAA
TCCCAGTCCGCACTGGGCACCGCTATCGAGCGTCTGTCTTCTGGTCTGCGTATCAACAGC
GCGAAAGACGATGCGGCAGGTCAGGCGATTGCTAACCGTTTCACCGCGAACATCAAAGGT
CTGACTCAGGCTTCCCGTAACGCTAACGACGGTATCTCCATTGCGCAGACCACTGAAGGC
GCGCTGAACGAAATCAACAACAACCTGCAGCGTGTGCGTGAACTGGCGGTTCAGTCTGCT
AACAGCACCAACTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATCACCCAGCGCCTG
AACGAAATCGACCGTGTATCCGGCCAGACTCAGTTCAACGGCGTGAAAGTCCTGGCGCAG
GACAACACCCTGACCATCCAGGTTGGCGCCAACGACGGTGAAACTATCGATATCGATCTG
AAGCAGATCAACTCTCAGACCCTGGGTCTGGACTCACTGAACGTGCAGAAAGCGTATGAT
GTGAAAGATACAGCAGTAACAACGAAAGCTTATGCCAATAATGGTACTACACTGGATGTA
TCGGGTCTTGATGATGCAGCTATTAAAGCGGCTACGGGTGGTACGAATGGTACGGCTTCT
GTAACCGGTGGTGCGGTTAAATTTGACGCAGATAATAACAAGTACTTTGTTACTATTGGT
GGCTTTACTGGTGCTGATGCCGCCAAAAATGGCGATTATGAAGTTAACGTTGCTACTGAC
GGTACAGTAACCCTTGCGGCTGGCGCAACTAAAACCACAATGCCTGCTGGTGCGACAACT
AAAACAGAAGTACAGGAGTTAAAAGATACACCGGCAGTTGTTTCAGCAGATGCTAAAAAT
GCCTTAATTGCTGGCGGCGTTGACGCTACCGATGCTAATGGCGCTGAGTTGGTCAAAATG
TCTTATACCGATAAAAATGGTAAGACAATTGAAGGCGGTTATGCGCTTAAAGCTGGCGAT
AAGTATTACGCCGCAGATTACGATGAAGCGACAGGAGCAATTAAAGCTAAAACCACAAGT
TATACTGCTGCTGACGGCACTACCAAAACAGCGGCTAACCAACTGGGTGGCGTAGACGGT
AAAACCGAAGTCGTTACTATCGACGGTAAAACCTACAATGCCAGCAAAGCCGCTGGTCAT
GATTTCAAAGCACAACCAGAGCTGGCGGAAGCAGCCGCTAAAACCACCGAAAACCCGCTG
CAGAAAATTGATGCCGCGCTGGCGCAGGTGGATGCGCTGCGCTCTGATCTGGGTGCGGTA
CAAAACCGTTTCAACTCTGCTATCACCAACCTGGGCAATACCGTAAACAATCTGTCTGAA
GCGCGTAGCCGTATCGAAGATTCCGACTACGCGACCGAAGTTTCCAACATGTCTCGCGCG
CAGATTTTGCAGCAGGCCGGTACTTCCGTTCTGGCGCAGGCTAACCAGGTCCCGCAGAAC
GTGCTGAGCCTGTTAGCGGGTGCGAAACCGCTGTCTCTGCGCGATTGCTCTGTGGCAGGC
TGGCTGCTGGGCAACCCTATGTGTGACGAATTCATTAACGTTCCGGAATGGTCTGATATT
GCCGAAAAAGCTAACCCTGTCAACGATCTGTGTTACCCTGGTGACTTTAACGATTACGAA
GAACTGAAGCACCTGCTGTCTCGTATCAATCACTTCGAGAAAATCCAGATCATCCCGAAA
TCCTCCTGGAGCTCCCACGAAGCTTCTCTGGGCGTATCCTCCGCGTGCCCGTACCAGGGC
AAATCCTCTTTCTTTCGTAACGTTGTTTGGCTGATCAAGAAAAACTCCACCTACCCGACG
ATCAAGCGTAGCTATAATAACACCAACCAGGAAGACCTGCTGGTTCTGTGGGGCATCCAC
CATCCAAACGATGCTGCGGAACAGACCAAGCTGTACCAGAACCCGACCACCTACATCAGC
GTGGGCACCTCTACGCTGAACCAGCGTCTGGTACCGCGTATCGCAACCCGCAGCAAGGTA
AACGGTCAAAGCGGCCGCATGGAATTTTTCTGGACCATCCTGAAACCGAACGACGCAATC
AACTTCGAATCTAACGGCAATTTCATCGCTCCGGAGTATGCGTACAAAATCGTAAAGAAA
GGTGATAGCACTATCATGAAATCCGAGTAGTAA

SEQ ID NO: 72 HA1-2mut VN

GGTGCGAAACCGCTGTCTCTGCGCGATTGCTCTGTGGCAGGCTGGCTGCTGGGCAACCCT
ATGTGTGACGAATTCATTAACGTTCCGGAATGGTCTGATATTGCCGAAAAAGCTAACCCT
GTCAACGATCTGTGTTACCCTGGTGACTTTAACGATTACGAAGAACTGAAGCACCTGCTG
TCTCGTATCAATCACTTCGAGAAAATCCAGATCATCCCGAAATCCTCCTGGAGCTCCCAC
GAAGCTTCTCTGGGCGTATCCTCCGCGTGCCCGTACCAGGGCAAATCCTCTTTCTTTCGT
AACGTTGTTTGGCTGATCAAGAAAAACTCCACCTACCCGACGATCAAGCGTAGCTATAAT
AACACCAACCAGGAAGACCTGCTGGTTCTGTGGGGCATCCACCATCCAAACGATGCTGCG
GAACAGACCAAGCTGTACCAGAACCCGACCACCTACATCAGCGTGGGCACCTCTACGCTG
AACCAGCGTCTGGTACCGCGTATCGCAACCCGCAGCAAGGTAAACGGTCAAAGCGGCCGC
ATGGAATTTTTCTGGACCATCCTGAAACCGAACGACGCAATCAACTTCGAATCTAACGGC
AATTTCATCGCTCCGGAGTATGCGTACAAAATCGTAAAGAAAGGTGATAGCACTATCATG
AAATCCGAGTAGTAA

SEQ ID NO: 73 STF2.HA1-1 IND

ATGGCACAAGTAATCAACACTAACAGTCTGTCGCTGCTGACCCAGAATAACCTGAACAAA
TCCCAGTCCGCACTGGGCACCGCTATCGAGCGTCTGTCTTCTGGTCTGCGTATCAACAGC
GCGAAAGACGATGCGGCAGGTCAGGCGATTGCTAACCGTTTCACCGCGAACATCAAAGGT
CTGACTCAGGCTTCCCGTAACGCTAACGACGGTATCTCCATTGCGCAGACCACTGAAGGC
GCGCTGAACGAAATCAACAACAACCTGCAGCGTGTGCGTGAACTGGCGGTTCAGTCTGCT
AACAGCACCAACTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATCACCCAGCGCCTG
AACGAAATCGACCGTGTATCCGGCCAGACTCAGTTCAACGGCGTGAAAGTCCTGGCGCAG
GACAACACCCTGACCATCCAGGTTGGCGCCAACGACGGTGAAACTATCGATATCGATCTG
AAGCAGATCAACTCTCAGACCCTGGGTCTGGACTCACTGAACGTGCAGAAAGCGTATGAT
GTGAAAGATACAGCAGTAACAACGAAAGCTTATGCCAATAATGGTACTACACTGGATGTA
TCGGGTCTTGATGATGCAGCTATTAAAGCGGCTACGGGTGGTACGAATGGTACGGCTTCT
GTAACCGGTGGTGCGGTTAAATTTGACGCAGATAATAACAAGTACTTTGTTACTATTGGT
GGCTTTACTGGTGCTGATGCCGCCAAAAATGGCGATTATGAAGTTAACGTTGCTACTGAC
GGTACAGTAACCCTTGCGGCTGGCGCAACTAAAACCACAATGCCTGCTGGTGCGACAACT
AAAACAGAAGTACAGGAGTTAAAAGATACACCGGCAGTTGTTTCAGCAGATGCTAAAAAT
GCCTTAATTGCTGGCGGCGTTGACGCTACCGATGCTAATGGCGCTGAGTTGGTCAAAATG
TCTTATACCGATAAAAATGGTAAGACAATTGAAGGCGGTTATGCGCTTAAAGCTGGCGAT
AAGTATTACGCCGCAGATTACGATGAAGCGACAGGAGCAATTAAAGCTAAAACCACAAGT
TATACTGCTGCTGACGGCACTACCAAAACAGCGGCTAACCAACTGGGTGGCGTAGACGGT
AAAACCGAAGTCGTTACTATCGACGGTAAAACCTACAATGCCAGCAAAGCCGCTGGTCAT
GATTTCAAAGCACAACCAGAGCTGGCGGAAGCAGCCGCTAAAACCACCGAAAACCCGCTG
CAGAAAATTGATGCCGCGCTGGCGCAGGTGGATGCGCTGCGCTCTGATCTGGGTGCGGTA
CAAAACCGTTTCAACTCTGCTATCACCAACCTGGGCAATACCGTAAACAATCTGTCTGAA
GCGCGTAGCCGTATCGAAGATTCCGACTACGCGACCGAAGTTTCCAACATGTCTCGCGCG
CAGATTTTGCAGCAGGCCGGTACTTCCGTTCTGGCGCAGGCTAACCAGGTCCCGCAGAAC
GTGCTGTCTCTGTTAGCGGAGAAAACGCATAACGGTAAACTGTGCGACCTGGACGGCGTG
AAGCCGCTGATCCTGCGTGACTGTTCTGTTGCTGGCTGGCTGCTGGGCAACCCGATGTGC
GATGAGTTCATCAACGTTCCGGAGTGGTCTTATATTGTTGAAAAAGCGAACCCGACCAAT
GACCTGTGCTACCCGGGCTCTTTTAATGACTACGAGGAGCTGAAACATCTGCTGTCTCGC
ATCAACCATTTCGAAAAAATTCAGATCATCCCGAAATCTTCCTGGAGCGATCACGAAGCT
TCTTCCGGCGTATCTTCTGCATGCCCTTACCTGGGTTCCCCGTCTTTCTTCCGTAACGTG
GTTTGGCTGATCAAGAAGAACTCTACGTACCCGACCATCAAGAAATCCTATAACAACACT
AACCAGGAAGATCTGCTGGTGCTGTGGGGTATCCATCATCCAAACGATGCAGCCGAACAG
ACCCGCCTGTACCAGAATCCGACCACGTACATCTCTATCGGCACTTCTACCCTGAATCAA
CGCCTGGTGCCGAAAATCGCAACTCGCAGCAAAGTCAACGGCCAATCTGGTCGCATGGAG
TTTTTCTGGACCATCCTGAAACCGAACGACGCGATTAACTTCGAAAGCAACGGCAACTTC
ATTGCACCGGAATATGCTTACAAGATCGTTAAAAAGGGTGATTCTGCGATCATGAAAAGC
GAGCTGGAATACGGCAACTGCAACACTAAATGCCAGACCCCGATGGGTGCAATCAATTCC
AGCATGCCATTTCACAACATCCACCCGCTGACCATCGGCGAGTGTCCAAAATACGTTAAA
TAATAG

SEQ ID NO: 74 FOR PRIMER

AGTGGCTGAGCCTGTTAGCGGAGAAAACGCATAACGGTAAACTGTG

SEQ ID NO: 75 HA0s IND

GCTGAGCCTGTTACGTGATCAAATCTGCATCGGTTACCACGCAAACAACTCCACTGAACA
AGTGGATACGATCATGGAAAAGAACGTGACCGTGACCCACGCTCAAGATATCCTGGAGAA
AACGCATAACGGTAAACTGTGCGACCTGGACGGCGTGAAGCCGCTGATCCTGCGTGACTG
TTCTGTTGCTGGCTGGCTGCTGGGCAACCCGATGTGCGATGAGTTCATCAACGTTCCGGA
GTGGTCTTATATTGTTGAAAAAGCGAACCCGACCAATGACCTGTGCTACCCGGGCTCTTT
TAATGACTACGAGGAGCTGAAACATCTGCTGTCTCGCATCAACCATTTCGAAAAAATTCA

GATCATCCCGAAATCTTCCTGGAGCGATCACGAAGCTTCTTCCGGCGTATCTTCTGCATG
CCCTTACCTGGGTTCCCCGTCTTTCTTCCGTAACGTGGTTTGGCTGATCAAGAAGAACTC
TACGTACCCGACCATCAAGAAATCCTATAACAACACTAACCAGGAAGATCTGCTGGTGCT
GTGGGGTATCCATCATCCAAACGATGCAGCCGAACAGACCCGCCTGTACCAGAATCCGAC
CACGTACATCTCTATCGGCACTTCTACCCTGAATCAACGCCTGGTGCCGAAAATCGCAAC
TCGCAGCAAAGTCAACGGCCAATCTGGTCGCATGGAGTTTTTCTGGACCATCCTGAAACC
GAACGACGCGATTAACTTCGAAAGCAACGGCAACTTCATTGCACCGGAATATGCTTACAA
GATCGTTAAAAAGGGTGATTCTGCGATCATGAAAAGCGAGCTGGAATACGGCAACTGCAA
CACTAAATGCCAGACCCCGATGGGTGCAATCAATTCCAGCATGCCATTTCACAACATCCA
CCCGCTGACCATCGGCGAGTGTCCAAAATACGTTAAATCCAATCGTCTGGTGCTGGCAAC
GGGCCTGCGCAACTCTCCGCAACGTGAGTCTCGTCGCAAAAAGCGTGGCCTGTTTGGCGC
AATTGCGGGCTTCATCGAGGGCGGCTGGCAGGGTATGGTTGATGGTTGGTACGGTTATCA
CCATTCCAACGAGCAGGGCTCTGGTTACGCTGCGGACAAAGAGTCCACGCAGAAAGCGAT
CGATGGCGTCACCAACAAAGTGAACTCTATCATCGACAAGATGAACACTCAGTTCGAGGC
TGTGGGCCGTGAATTCAACAATCTGGAACGTCGCATCGAAAACCTGAACAAAAAGATGGA
AGACGGTTTCCTGGACGTATGGACCTATAACGCAGAGCTGCTGGTTCTGATGGAAAACGA
ACGTACTCTGGACTTCCACGATTCTAACGTTAAAAACCTGTACGACAAAGTTCGTCTGCA
GCTGCGCGATAATGCAAAAGAACTGGGTAACGGCTGCTTTGAATTTTACCACAAATGTGA
CAACGAATGCATGGAAAGCATCCGTAACGGTACCTACAATTACCCACAGTACTCCGAAGA
AGCGCGTCTGAAACGTGAAGAAATCTCCGGTGTAAAACTGGAATCCATTGGCACGTATCA
GATTCTGTCTATCTACTCCACGGTAGCGTCCTCTCTGGCGCTGGCAATTATGATGGCCGG
CCTGTCTCTGTGGATGTGCTCTAACGGCTCTCTGCAGTGCCGCATCTGCATCTAATAGGC
TGAGC

SEQ ID NO: 76 STF2.HA1-2 IND

ATGGCACAAGTAATCAACACTAACAGTCTGTCGCTGCTGACCCAGAATAACCTGAACAAA
TCCCAGTCCGCACTGGGCACCGCTATCGAGCGTCTGTCTTCTGGTCTGCGTATCAACAGC
GCGAAAGACGATGCGGCAGGTCAGGCGATTGCTAACCGTTTCACCGCGAACATCAAAGGT
CTGACTCAGGCTTCCCGTAACGCTAACGACGGTATCTCCATTGCGCAGACCACTGAAGGC
GCGCTGAACGAAATCAACAACAACCTGCAGCGTGTGCGTGAACTGGCGGTTCAGTCTGCT
AACAGCACCAACTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATCACCCAGCGCCTG
AACGAAATCGACCGTGTATCCGGCCAGACTCAGTTCAACGGCGTGAAAGTCCTGGCGCAG
GACAACACCCTGACCATCCAGGTTGGCGCCAACGACGGTGAAACTATCGATATCGATCTG
AAGCAGATCAACTCTCAGACCCTGGGTCTGGACTCACTGAACGTGCAGAAAGCGTATGAT
GTGAAAGATACAGCAGTAACAACGAAAGCTTATGCCAATAATGGTACTACACTGGATGTA
TCGGGTCTTGATGATGCAGCTATTAAAGCGGCTACGGGTGGTACGAATGGTACGGCTTCT
GTAACCGGTGGTGCGGTTAAATTTGACGCAGATAATAACAAGTACTTTGTTACTATTGGT
GGCTTTACTGGTGCTGATGCCGCCAAAAATGGCGATTATGAAGTTAACGTTGCTACTGAC
GGTACAGTAACCCTTGCGGCTGGCGCAACTAAAACCACAATGCCTGCTGGTGCGACAACT
AAAACAGAAGTACAGGAGTTAAAAGATACACCGGCAGTTGTTTCAGCAGATGCTAAAAAT
GCCTTAATTGCTGGCGGCGTTGACGCTACCGATGCTAATGGCGCTGAGTTGGTCAAAATG
TCTTATACCGATAAAAATGGTAAGACAATTGAAGGCGGTTATGCGCTTAAAGCTGGCGAT
AAGTATTACGCCGCAGATTACGATGAAGCGACAGGAGCAATTAAAGCTAAAACCACAAGT
TATACTGCTGCTGACGGCACTACCAAAACAGCGGCTAACCAACTGGGTGGCGTAGACGGT
AAAACCGAAGTCGTTACTATCGACGGTAAAACCTACAATGCCAGCAAAGCCGCTGGTCAT
GATTTCAAAGCACAACCAGAGCTGGCGGAAGCAGCCGCTAAAACCACCGAAAACCCGCTG
CAGAAAATTGATGCCGCGCTGGCGCAGGTGGATGCGCTGCGCTCTGATCTGGGTGCGGTA
CAAAACCGTTTCAACTCTGCTATCACCAACCTGGGCAATACCGTAAACAATCTGTCTGAA
GCGCGTAGCCGTATCGAAGATTCCGACTACGCGACCGAAGTTTCCAACATGTCTCGCGCG
CAGATTTTGCAGCAGGCCGGTACTTCCGTTCTGGCGCAGGCTAACCAGGTCCCGCAGAAC
GTGCTGAGCCTGTTAGCAGGTGTGAAACCGCTGATTCTGCGCGATTGTTCTGTTGCTGGC
TGGCTGCTGGGCAACCCGATGTGCGATGAGTTCATCAACGTTCCGGAGTGGTCTTATATT
GTTGAAAAAGCGAACCCGACCAATGACCTGTGCTACCCGGGCTCTTTTAATGACTACGAG

GAGCTGAAACATCTGCTGTCTCGCATCAACCATTTCGAAAAAATTCAGATCATCCCGAAA
TCTTCCTGGAGCGATCACGAAGCTTCTTCCGGCGTATCTTCTGCATGCCCTTACCTGGGT
TCCCCGTCTTTCTTCCGTAACGTGGTTTGGCTGATCAAGAAGAACTCTACGTACCCGACC
ATCAAGAAATCCTATAACAACACTAACCAGGAAGATCTGCTGGTGCTGTGGGGTATCCAT
CATCCAAACGATGCAGCCGAACAGACCCGCCTGTACCAGAATCCGACCACGTACATCTCT
ATCGGCACTTCTACCCTGAATCAACGCCTGGTGCCGAAAATCGCAACTCGCAGCAAAGTC
AACGGCCAATCTGGTCGCATGGAGTTTTTCTGGACCATCCTGAAACCGAACGACGCGATT
AACTTCGAAAGCAACGGCAACTTCATTGCACCGGAATATGCTTACAAGATCGTTAAAAAG
GGTGATTCTGCGATCATGAAAAGCGAGTAATAG

SEQ ID NO: 77 REV PRIMER
AGTCGCTCAGCCTATTACTCGCTTTTCATGATCGCAGAATCAC

SEQ ID NO: 78 STF2.HA1-2mut IND
ATGGCACAAGTAATCAACACTAACAGTCTGTCGCTGCTGACCCAGAATAACCTGAACAAA
TCCCAGTCCGCACTGGGCACCGCTATCGAGCGTCTGTCTTCTGGTCTGCGTATCAACAGC
GCGAAAGACGATGCGGCAGGTCAGGCGATTGCTAACCGTTTCACCGCGAACATCAAAGGT
CTGACTCAGGCTTCCCGTAACGCTAACGACGGTATCTCCATTGCGCAGACCACTGAAGGC
GCGCTGAACGAAATCAACAACAACCTGCAGCGTGTGCGTGAACTGGCGGTTCAGTCTGCT
AACAGCACCAACTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATCACCCAGCGCCTG
AACGAAATCGACCGTGTATCCGGCCAGACTCAGTTCAACGGCGTGAAAGTCCTGGCGCAG
GACAACACCCTGACCATCCAGGTTGGCGCCAACGACGGTGAAACTATCGATATCGATCTG
AAGCAGATCAACTCTCAGACCCTGGGTCTGGACTCACTGAACGTGCAGAAAGCGTATGAT
GTGAAAGATACAGCAGTAACAACGAAAGCTTATGCCAATAATGGTACTACACTGGATGTA
TCGGGTCTTGATGATGCAGCTATTAAAGCGGCTACGGGTGGTACGAATGGTACGGCTTCT
GTAACCGGTGGTGCGGTTAAATTTGACGCAGATAATAACAAGTACTTTGTTACTATTGGT
GGCTTTACTGGTGCTGATGCCGCCAAAAATGGCGATTATGAAGTTAACGTTGCTACTGAC
GGTACAGTAACCCTTGCGGCTGGCGCAACTAAAACCACAATGCCTGCTGGTGCGACAACT
AAAACAGAAGTACAGGAGTTAAAAGATACACCGGCAGTTGTTTCAGCAGATGCTAAAAAT
GCCTTAATTGCTGGCGGCGTTGACGCTACCGATGCTAATGGCGCTGAGTTGGTCAAAATG
TCTTATACCGATAAAAATGGTAAGACAATTGAAGGCGGTTATGCGCTTAAAGCTGGCGAT
AAGTATTACGCCGCAGATTACGATGAAGCGACAGGAGCAATTAAAGCTAAAACCACAAGT
TATACTGCTGCTGACGGCACTACCAAAACAGCGGCTAACCAACTGGGTGGCGTAGACGGT
AAAACCGAAGTCGTTACTATCGACGGTAAAACCTACAATGCCAGCAAAGCCGCTGGTCAT
GATTTCAAAGCACAACCAGAGCTGGCGGAAGCAGCCGCTAAAACCACCGAAAACCCGCTG
CAGAAAATTGATGCCGCGCTGGCGCAGGTGGATGCGCTGCGCTCTGATCTGGGTGCGGTA
CAAAACCGTTTCAACTCTGCTATCACCAACCTGGGCAATACCGTAAACAATCTGTCTGAA
GCGCGTAGCCGTATCGAAGATTCCGACTACGCGACCGAAGTTTCCAACATGTCTCGCGCG
CAGATTTTGCAGCAGGCCGGTACTTCCGTTCTGGCGCAGGCTAACCAGGTCCCGCAGAAC
GTGCTGTCTCTGTTAGCGGGCGCAAAGCCGCTGTCTCTGCGTGACTGTTCTGTTGCTGGC
TGGCTGCTGGGCAACCCGATGTGCGATGAGTTCATCAACGTTCCGGAGTGGTCTGATATT
GCGGAAAAAGCGAACCCGACCAATGACCTGTGCTACCCGGGCTCTTTTAATGACTACGAG
GAGCTGAAACATCTGCTGTCTCGCATCAACCATTTCGAAAAAATTCAGATCATCCCGAAA
TCTTCCTGGAGCGATCACGAAGCTTCTTCCGGCGTATCTTCTGCATGCCCTTACCTGGGT
TCCCCGTCTTTCTTCCGTAACGTGGTTTGGCTGATCAAGAAGAACTCTACGTACCCGACC
ATCAAGAAATCCTATAACAACACTAACCAGGAAGATCTGCTGGTGCTGTGGGGTATCCAT
CATCCAAACGATGCAGCCGAACAGACCCGCCTGTACCAGAATCCGACCACGTACATCTCT
ATCGGCACTTCTACCCTGAATCAACGCCTGGTGCCGAAAATCGCAACTCGCAGCAAAGTC

AACGGCCAATCTGGTCGCATGGAGTTTTTCTGGACCATCCTGAAACCGAACGACGCGATT
AACTTCGAAAGCAACGGCAACTTCATTGCACCGGAATATGCTTACAAGATCGTTAAAAAG
GGTGATTCTGCGATCATGAAAAGCGAGTAATAG

SEQ ID 79 HA1-2mut IND

GGCGCAAAGCCGCTGTCTCTGCGTGACTGTTCTGTTGCTGGCTGGCTGCTGGGCAACCCG
ATGTGCGATGAGTTCATCAACGTTCCGGAGTGGTCTGATATTGCGGAAAAAGCGAACCCG
ACCAATGACCTGTGCTACCCGGGCTCTTTTAATGACTACGAGGAGCTGAAACATCTGCTG
TCTCGCATCAACCATTTCGAAAAAATTCAGATCATCCCGAAATCTTCCTGGAGCGATCAC
GAAGCTTCTTCCGGCGTATCTTCTGCATGCCCTTACCTGGGTTCCCCGTCTTTCTTCCGT
AACGTGGTTTGGCTGATCAAGAAGAACTCTACGTACCCGACCATCAAGAAATCCTATAAC
AACACTAACCAGGAAGATCTGCTGGTGCTGTGGGGTATCCATCATCCAAACGATGCAGCC
GAACAGACCCGCCTGTACCAGAATCCGACCACGTACATCTCTATCGGCACTTCTACCCTG
AATCAACGCCTGGTGCCGAAAATCGCAACTCGCAGCAAAGTCAACGGCCAATCTGGTCGC
ATGGAGTTTTTCTGGACCATCCTGAAACCGAACGACGCGATTAACTTCGAAAGCAACGGC
AACTTCATTGCACCGGAATATGCTTACAAGATCGTTAAAAAGGGTGATTCTGCGATCATG
AAAAGCGAGTAATAGGCTGAGC

SEQ ID NO: 80 STF2.HA1-1 NC

ATGGCACAAGTAATCAACACTAACAGTCTGTCGCTGCTGACCCAGAATAACCTGAACAAA
TCCCAGTCCGCACTGGGCACCGCTATCGAGCGTCTGTCTTCTGGTCTGCGTATCAACAGC
GCGAAAGACGATGCGGCAGGTCAGGCGATTGCTAACCGTTTCACCGCGAACATCAAAGGT
CTGACTCAGGCTTCCCGTAACGCTAACGACGGTATCTCCATTGCGCAGACCACTGAAGGC
GCGCTGAACGAAATCAACAACAACCTGCAGCGTGTGCGTGAACTGGCGGTTCAGTCTGCT
AACAGCACCAACTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATCACCCAGCGCCTG
AACGAAATCGACCGTGTATCCGGCCAGACTCAGTTCAACGGCGTGAAAGTCCTGGCGCAG
GACAACACCCTGACCATCCAGGTTGGCGCCAACGACGGTGAAACTATCGATATCGATCTG
AAGCAGATCAACTCTCAGACCCTGGGTCTGGACTCACTGAACGTGCAGAAAGCGTATGAT
GTGAAAGATACAGCAGTAACAACGAAAGCTTATGCCAATAATGGTACTACACTGGATGTA
TCGGGTCTTGATGATGCAGCTATTAAAGCGGCTACGGGTGGTACGAATGGTACGGCTTCT
GTAACCGGTGGTGCGGTTAAATTTGACGCAGATAATAACAAGTACTTTGTTACTATTGGT
GGCTTTACTGGTGCTGATGCCGCCAAAAATGGCGATTATGAAGTTAACGTTGCTACTGAC
GGTACAGTAACCCTTGCGGCTGGCGCAACTAAAACCACAATGCCTGCTGGTGCGACAACT
AAAACAGAAGTACAGGAGTTAAAAGATACACCGGCAGTTGTTTCAGCAGATGCTAAAAAT
GCCTTAATTGCTGGCGGCGTTGACGCTACCGATGCTAATGGCGCTGAGTTGGTCAAAATG
TCTTATACCGATAAAAATGGTAAGACAATTGAAGGCGGTTATGCGCTTAAAGCTGGCGAT
AAGTATTACGCCGCAGATTACGATGAAGCGACAGGAGCAATTAAAGCTAAAACCACAAGT
TATACTGCTGCTGACGGCACTACCAAAACAGCGGCTAACCAACTGGGTGGCGTAGACGGT
AAAACCGAAGTCGTTACTATCGACGGTAAAACCTACAATGCCAGCAAAGCCGCTGGTCAT
GATTTCAAAGCACAACCAGAGCTGGCGGAAGCAGCCGCTAAAACCACCGAAAACCCGCTG
CAGAAAATTGATGCCGCGCTGGCGCAGGTGGATGCGCTGCGCTCTGATCTGGGTGCGGTA
CAAAACCGTTTCAACTCTGCTATCACCAACCTGGGCAATACCGTAAACAATCTGTCTGAA
GCGCGTAGCCGTATCGAAGATTCCGACTACGCGACCGAAGTTTCCAACATGTCTCGCGCG
CAGATTTTGCAGCAGGCCGGTACTTCCGTTCTGGCGCAGGCTAACCAGGTCCCGCAGAAC
GTGCTGAGCCTGTTAGCGTCCCACAACGGTAAACTGTGTCTGCTGAAAGGCATCGCACCG
CTGCAGCTGGGTAACTGTAGCGTTGCAGGTTGGATCCTGGGTAACCCGGAGTGCGAACTG
CTGATTTCCAAGGAGAGCTGGTCTTACATTGTCGAAACGCCGAATCCGGAAAACGGCACT
TGTTATCCGGGTTACTTCGCCGATTACGAAGAACTGCGTGAACAACTGTCTTCCGTGAGC
TCTTTCGAACGTTTCGAGATCTTCCCGAAAGAGAGCAGCTGGCCAAACCACACTGTTACC
GGTGTGTCTGCGAGCTGTTCTCACAACGGCAAGTCTTCTTTCTACCGCAACCTGCTGTGG
CTGACGGGTAAGAATGGCCTGTATCCGAACCTGTCTAAATCTTACGTAAACAACAAAGAG
AAAGAGGTGCTGGTCCTGTGGGGCGTACACCATCCACCAAATATCGGCAACCAGCGCGCC
CTGTACCACACCGAAAACGCTTATGTGTCCGTGGTGAGCTCCCATTACAGCCGTCGTTTT

ACTCCGGAGATTGCCAAACGTCCGAAAGTTCGTGATCAGGAAGGCCGTATTAACTACTAC
TGGACTCTGCTGGAGCCGGGCGATACCATCATTTTCGAGGCAAACGGCAACCTGATTGCG
CCATGGTACGCGTTCGCCCTGAGCCGTGGTTTTGGCTCCGGTATTATCACCTCTAACGCG
CCAATGGACGAATGCGACGCGAAATGCCAAACGCCGCAGGGCGCAATCAACAGCAGCCTG
CCGTTCCAGAACGTTCACCCGGTTACCATCGGCGAATGCCCTAAATACGTGCGCTAATAG

SEQ ID NO: 81 FOR PRIMER
AACCAGGTCCCGCAGAACGTGCTGAGCCTGTTAGCGTCCCACAACGGTAAACTGTGTCTG
CTGAAAGGCATCGCACCGCTGCAG

SEQ ID NO: 82 REV PRIMER
GAACAGCTCAGTTCCTACGAGCTCAGCCTATTAGCGCACGTATTTAGGGCATTCGCCGATG
GTAACCGGGTGAACGTT

SEQ ID NO: 83 HA0s NC
GACACGATCTGTATTGGTTATCATGCAAACAACTCTACTGACACTGTAGATACTGTGCTG
GAAAAGAACGTAACCGTTACCCACAGCGTTAACCTGCTGGAAGATTCCCACAACGGTAAA
CTGTGTCTGCTGAAAGGCATCGCACCGCTGCAGCTGGGTAACTGTAGCGTTGCAGGTTGG
ATCCTGGGTAACCCGGAGTGCGAACTGCTGATTTCCAAGGAGAGCTGGTCTTACATTGTC
GAAACGCCGAATCCGGAAAACGGCACTTGTTATCCGGGTTACTTCGCCGATTACGAAGAA
CTGCGTGAACAACTGTCTTCCGTGAGCTCTTTCGAACGTTTCGAGATCTTCCCGAAAGAG
AGCAGCTGGCCAAACCACACTGTTACCGGTGTGTCTGCGAGCTGTTCTCACAACGGCAAG
TCTTCTTTCTACCGCAACCTGCTGTGGCTGACGGGTAAGAATGGCCTGTATCCGAACCTG
TCTAAATCTTACGTAAACAACAAAGAGAAAGAGGTGCTGGTCCTGTGGGGCGTACACCAT
CCACCAAATATCGGCAACCAGCGCGCCCTGTACCACACCGAAAACGCTTATGTGTCCGTG
GTGAGCTCCCATTACAGCCGTCGTTTTACTCCGGAGATTGCCAAACGTCCGAAAGTTCGT
GATCAGGAAGGCCGTATTAACTACTACTGGACTCTGCTGGAGCCGGGCGATACCATCATT
TTCGAGGCAAACGGCAACCTGATTGCGCCATGGTACGCGTTCGCCCTGAGCCGTGGTTTT
GGCTCCGGTATTATCACCTCTAACGCGCCAATGGACGAATGCGACGCGAAATGCCAAACG
CCGCAGGGCGCAATCAACAGCAGCCTGCCGTTCCAGAACGTTCACCCGGTTACCATCGGC
GAATGCCCTAAATACGTGCGCTCCGCCAAACTGCGCATGGTTACTGGTCTGCGTAACATC
CCGAGCATTCAGTCTCGCGGTCTGTTCGGTGCGATCGCGGGCTTCATTGAAGGCGGTTGG
ACCGGTATGGTGGATGGTTGGTACGGCTACCATCACCAGAACGAACAGGGTAGCGGTTAC
GCTGCCGACCAGAAATCCACCCAGAACGCTATTAACGGTATCACCAACAAAGTTAACAGC
GTAATTGAGAAGATGAACACGCAGTTCACCGCCGTAGGTAAGGAATTCAACAAGCTGGAA
CGTCGCATGGAAAACCTGAACAAAAAGGTGGACGACGGCTTTCTGGACATCTGGACCTAC
AACGCTGAACTGCTGGTGCTGCTGGAAAACGAACGTACCCTGGATTTCCACGACTCTAAT
GTTAAAAACCTGTACGAAAAGGTCAAGTCTCAACTGAAAAACAATGCGAAGGAAATCGGC
AACGGCTGTTTCGAATTCTACCATAAATGCAACAACGAATGCATGGAATCCGTTAAAAAC
GGTACCTATGACTACCCTAAATACTCCGAAGAAAGCAAACTGAACCGCGAGAAAATCGAT
GGTGTAAAACTGGAATCTATGGGTGTTTACCAGATCCTGGCGATCTACTCCACGGTAGCC
AGCAGCCTGGTTCTGCTGGTTAGCCTGGGTGCAATTAGCTTCTGGATGTGCTCTAACGGC
AGCCTGCAATGCCGCATCTGTAATAGGCTGAGC

SEQ ID NO: 84 STF2HA1-2 NC

ATGGCACAAGTAATCAACACTAACAGTCTGTCGCTGCTGACCCAGAATAACCTGAACAAA
TCCCAGTCCGCACTGGGCACCGCTATCGAGCGTCTGTCTTCTGGTCTGCGTATCAACAGC
GCGAAAGACGATGCGGCAGGTCAGGCGATTGCTAACCGTTTCACCGCGAACATCAAAGGT
CTGACTCAGGCTTCCCGTAACGCTAACGACGGTATCTCCATTGCGCAGACCACTGAAGGC
GCGCTGAACGAAATCAACAACAACCTGCAGCGTGTGCGTGAACTGGCGGTTCAGTCTGCT
AACAGCACCAACTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATCACCCAGCGCCTG
AACGAAATCGACCGTGTATCCGGCCAGACTCAGTTCAACGGCGTGAAAGTCCTGGCGCAG
GACAACACCCTGACCATCCAGGTTGGCGCCAACGACGGTGAAACTATCGATATCGATCTG
AAGCAGATCAACTCTCAGACCCTGGGTCTGGACTCACTGAACGTGCAGAAAGCGTATGAT
GTGAAAGATACAGCAGTAACAACGAAAGCTTATGCCAATAATGGTACTACACTGGATGTA
TCGGGTCTTGATGATGCAGCTATTAAAGCGGCTACGGGTGGTACGAATGGTACGGCTTCT
GTAACCGGTGGTGCGGTTAAATTTGACGCAGATAATAACAAGTACTTTGTTACTATTGGT
GGCTTTACTGGTGCTGATGCCGCCAAAAATGGCGATTATGAAGTTAACGTTGCTACTGAC
GGTACAGTAACCCTTGCGGCTGGCGCAACTAAAACCACAATGCCTGCTGGTGCGACAACT
AAAACAGAAGTACAGGAGTTAAAAGATACACCGGCAGTTGTTTCAGCAGATGCTAAAAAT
GCCTTAATTGCTGGCGGCGTTGACGCTACCGATGCTAATGGCGCTGAGTTGGTCAAAATG
TCTTATACCGATAAAAATGGTAAGACAATTGAAGGCGGTTATGCGCTTAAAGCTGGCGAT
AAGTATTACGCCGCAGATTACGATGAAGCGACAGGAGCAATTAAAGCTAAAACCACAAGT
TATACTGCTGCTGACGGCACTACCAAAACAGCGGCTAACCAACTGGGTGGCGTAGACGGT
AAAACCGAAGTCGTTACTATCGACGGTAAAACCTACAATGCCAGCAAAGCCGCTGGTCAT
GATTTCAAAGCACAACCAGAGCTGGCGGAAGCAGCCGCTAAAACCACCGAAAACCCGCTG
CAGAAAATTGATGCCGCGCTGGCGCAGGTGGATGCGCTGCGCTCTGATCTGGGTGCGGTA
CAAAACCGTTTCAACTCTGCTATCACCAACCTGGGCAATACCGTAAACAATCTGTCTGAA
GCGCGTAGCCGTATCGAAGATTCCGACTACGCGACCGAAGTTTCCAACATGTCTCGCGCG
CAGATTTTGCAGCAGGCCGGTACTTCCGTTCTGGCGCAGGCTAACCAGGTCCCGCAGAAC
GTGCTGAGCCTGTTAGCGAAAGGCATCGCACCGCTGCAGCTGGGTAACTGTAGCGTTGCA
GGTTGGATCCTGGGTAACCCGGAGTGCGAACTGCTGATTTCCAAGGAGAGCTGGTCTTAC
ATTGTCGAAACGCCGAATCCGGAAAACGGCACTTGTTATCCGGGTTACTTCGCCGATTAC
GAAGAACTGCGTGAACAACTGTCTTCCGTGAGCTCTTTCGAACGTTTCGAGATCTTCCCG
AAAGAGAGCAGCTGGCCAAACCACACTGTTACCGGTGTGTCTGCGAGCTGTTCTCACAAC
GGCAAGTCTTCTTTCTACCGCAACCTGCTGTGGCTGACGGGTAAGAATGGCCTGTATCCG
AACCTGTCTAAATCTTACGTAAACAACAAAGAGAAAGAGGTGCTGGTCCTGTGGGGCGTA
CACCATCCACCAAATATCGGCAACCAGCGCGCCCTGTACCACACCGAAAACGCTTATGTG
TCCGTGGTGAGCTCCCATTACAGCCGTCGTTTTACTCCGGAGATTGCCAAACGTCCGAAA
GTTCGTGATCAGGAAGGCCGTATTAACTACTACTGGACTCTGCTGGAGCCGGGCGATACC
ATCATTTTCGAGGCAAACGGCAACCTGATTGCGCCATGGTACGCGTTCGCCCTGAGCCGT
GGTTTTGGCTCCGGTATTATCACCTCTTAATAA

SEQ ID NO: 85 FOR PRIMER

AACCAGGTCCCGCAGAACGTGCTGAGCCTGTTAGCGAAAGGCATCGCACCGCTGCAGCTG
GGTAACTGTAGCGTTGCAGGTTGG

SEQ ID NO: 86 REV PRIMER

GAACAGCTCAGTTCCTACGAGCTCAGCCTATTAAGAGGTGATAATACCGGAGCCAAAACC
ACGGCTCAGGGCGAACGCGTACCA

SEQ ID NO: 87 STF2.HA1-2*mut* NC

ATGGCACAAGTAATCAACACTAACAGTCTGTCGCTGCTGACCCAGAATAACCTGAACAAA
TCCCAGTCCGCACTGGGCACCGCTATCGAGCGTCTGTCTTCTGGTCTGCGTATCAACAGC
GCGAAAGACGATGCGGCAGGTCAGGCGATTGCTAACCGTTTCACCGCGAACATCAAAGGT
CTGACTCAGGCTTCCCGTAACGCTAACGACGGTATCTCCATTGCGCAGACCACTGAAGGC
GCGCTGAACGAAATCAACAACAACCTGCAGCGTGTGCGTGAACTGGCGGTTCAGTCTGCT
AACAGCACCAACTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATCACCCAGCGCCTG
AACGAAATCGACCGTGTATCCGGCCAGACTCAGTTCAACGGCGTGAAAGTCCTGGCGCAG
GACAACACCCTGACCATCCAGGTTGGCGCCAACGACGGTGAAACTATCGATATCGATCTG
AAGCAGATCAACTCTCAGACCCTGGGTCTGGACTCACTGAACGTGCAGAAAGCGTATGAT
GTGAAAGATACAGCAGTAACAACGAAAGCTTATGCCAATAATGGTACTACACTGGATGTA
TCGGGTCTTGATGATGCAGCTATTAAAGCGGCTACGGGTGGTACGAATGGTACGGCTTCT
GTAACCGGTGGTGCGGTTAAATTTGACGCAGATAATAACAAGTACTTTGTTACTATTGGT
GGCTTTACTGGTGCTGATGCCGCCAAAAATGGCGATTATGAAGTTAACGTTGCTACTGAC
GGTACAGTAACCCTTGCGGCTGGCGCAACTAAAACCACAATGCCTGCTGGTGCGACAACT
AAAACAGAAGTACAGGAGTTAAAAGATACACCGGCAGTTGTTTCAGCAGATGCTAAAAAT
GCCTTAATTGCTGGCGGCGTTGACGCTACCGATGCTAATGGCGCTGAGTTGGTCAAAATG
TCTTATACCGATAAAAATGGTAAGACAATTGAAGGCGGTTATGCGCTTAAAGCTGGCGAT
AAGTATTACGCCGCAGATTACGATGAAGCGACAGGAGCAATTAAAGCTAAAACCACAAGT
TATACTGCTGCTGACGGCACTACCAAAACAGCGGCTAACCAACTGGGTGGCGTAGACGGT
AAAACCGAAGTCGTTACTATCGACGGTAAAACCTACAATGCCAGCAAAGCCGCTGGTCAT
GATTTCAAAGCACAACCAGAGCTGGCGGAAGCAGCCGCTAAAACCACCGAAAACCCGCTG
CAGAAAATTGATGCCGCGCTGGCGCAGGTGGATGCGCTGCGCTCTGATCTGGGTGCGGTA
CAAAACCGTTTCAACTCTGCTATCACCAACCTGGGCAATACCGTAAACAATCTGTCTGAA
GCGCGTAGCCGTATCGAAGATTCCGACTACGCGACCGAAGTTTCCAACATGTCTCGCGCG
CAGATTTTGCAGCAGGCCGGTACTTCCGTTCTGGCGCAGGCTAACCAGGTCCCGCAGAAC
GTGCTGAGCCTGTTAGCGAAAGGCGCGGCACCGCTGCAGCTGGGTAACTGTAGCGTTGCA
GGTTGGATCCTGGGTAACCCGGAGTGCGAACTGCTGATTTCCAAGGAGAGCTGGTCTGAT
ATTGCAGAAACGCCGAATCCGGAAAACGGCACTTGTTATCCGGGTTACTTCGCCGATTAC
GAAGAACTGCGTGAACAACTGTCTTCCGTGAGCTCTTTCGAACGTTTCGAGATCTTCCCG
AAAGAGAGCAGCTGGCCAAACCACACTGTTACCGGTGTGTCTGCGAGCTGTTCTCACAAC
GGCAAGTCTTCTTTCTACCGCAACCTGCTGTGGCTGACGGGTAAGAATGGCCTGTATCCG
AACCTGTCTAAATCTTACGTAAACAACAAAGAGAAAGAGGTGCTGGTCCTGTGGGGCGTA
CACCATCCACCAAATATCGGCAACCAGCGCGCCCTGTACCACACCGAAAACGCTTATGTG
TCCGTGGTGAGCTCCCATTACAGCCGTCGTTTTACTCCGGAGATTGCCAAACGTCCGAAA
GTTCGTGATCAGGAAGGCCGTATTAACTACTACTGGACTCTGCTGGAGCCGGGCGATACC
ATCATTTTCGAGGCAAACGGCAACCTGATTGCGCCATGGTACGCGTTCGCCCTGAGCCGT
GGTTTTGGCTCCGGTATTATCACCTCTTAATAG

SEQ ID NO: 88 HA1-2*mut* NC

GCTGAGCCTGTTAGCGAAAGGCGCGGCACCGCTGCAGCTGGGTAACTGTAGCGTTGCAGG
TTGGATCCTGGGTAACCCGGAGTGCGAACTGCTGATTTCCAAGGAGAGCTGGTCTGATAT
TGCAGAAACGCCGAATCCGGAAAACGGCACTTGTTATCCGGGTTACTTCGCCGATTACGA
AGAACTGCGTGAACAACTGTCTTCCGTGAGCTCTTTCGAACGTTTCGAGATCTTCCCGAA
AGAGAGCAGCTGGCCAAACCACACTGTTACCGGTGTGTCTGCGAGCTGTTCTCACAACGG
CAAGTCTTCTTTCTACCGCAACCTGCTGTGGCTGACGGGTAAGAATGGCCTGTATCCGAA
CCTGTCTAAATCTTACGTAAACAACAAAGAGAAAGAGGTGCTGGTCCTGTGGGGCGTACA
CCATCCACCAAATATCGGCAACCAGCGCGCCCTGTACCACACCGAAAACGCTTATGTGTC
CGTGGTGAGCTCCCATTACAGCCGTCGTTTTACTCCGGAGATTGCCAAACGTCCGAAAGT
TCGTGATCAGGAAGGCCGTATTAACTACTACTGGACTCTGCTGGAGCCGGGCGATACCAT
CATTTTCGAGGCAAACGGCAACCTGATTGCGCCATGGTACGCGTTCGCCCTGAGCCGTGG
TTTTGGCTCCGGTATTATCACCTCTTAATAGGCTGAGC

SEQ ID NO: 89 STF2.HA1-1his (PR8)

MAQVINTNSLSLLTQNNLNKSQSALGTAIERLSSGLRINSAKDDAAGQAIANRFTANIKGL
TQASRNANDGISIAQTTEGALNEINNNLQRVRELAVQSANSTNSQSDLDSIQAEITQRLNE
IDRVSGQTQFNGVKVLAQDNTLTIQVGANDGETIDIDLKQINSQTLGLDSLNVQKAYDVKD
TAVTTKAYANNGTTLDVSGLDDAAIKAATGGTNGTASVTGGAVKFDADNNKYFVTIGGFTG
ADAAKNGDYEVNVATDGTVTLAAGATKTTMPAGATTKTEVQELKDTPAVVSADAKNALIAG
GVDATDANGAELVKMSYTDKNGKTIEGGYALKAGDKYYAADYDEATGAIKAKTTSYTAADG
TTKTAANQLGGVDGKTEVVTIDGKTYNASKAAGHDFKAQPELAEAAAKTTENPLQKIDAAL
AQVDALRSDLGAVQNRFNSAITNLGNTVNNLSEARSRIEDSDYATEVSNMSRAQILQQAGT
SVLAQANQVPQNVLSLLASHNGKLCRLKGIAPLQLGKCNIAGWLLGNPECDPLLPVRSWSY
IVETPNSENGICYPGDFIDYEELREQLSSVSSFERFEIFPKESSWPNHNTNGVTAACSHEG
KSSFYRNLLWLTEKEGSYPKLKNSYVNKKGKEVLVLWGIHHPPNSKEQQNLYQNENAYVSV
VTSNYNRRFTPEIAERPKVRDQAGRMNYYWTLLKPGDTIIFEANGNLIAPMYAFALSRGFG
SGIITSNASMHECNTKCQTPLGAINSSLPYQNIHPVTIGECPKYVRHHHHHH

SEQ ID NO: 90 STF2.HA1-2(PR8)

MAQVINTNSLSLLTQNNLNKSQSALGTAIERLSSGLRINSAKDDAAGQAIANRFTANIKGL
TQASRNANDGISIAQTTEGALNEINNNLQRVRELAVQSANSTNSQSDLDSIQAEITQRLNE
IDRVSGQTQFNGVKVLAQDNTLTIQVGANDGETIDIDLKQINSQTLGLDSLNVQKAYDVKD
TAVTTKAYANNGTTLDVSGLDDAAIKAATGGTNGTASVTGGAVKFDADNNKYFVTIGGFTG
ADAAKNGDYEVNVATDGTVTLAAGATKTTMPAGATTKTEVQELKDTPAVVSADAKNALIAG
GVDATDANGAELVKMSYTDKNGKTIEGGYALKAGDKYYAADYDEATGAIKAKTTSYTAADG
TTKTAANQLGGVDGKTEVVTIDGKTYNASKAAGHDFKAQPELAEAAAKTTENPLQKIDAAL
AQVDALRSDLGAVQNRFNSAITNLGNTVNNLSEARSRIEDSDYATEVSNMSRAQILQQAGT
SVLAQANQVPQNVLSLLAKGIAPLQLGKCNIAGWLLGNPECDPLLPVRSWSYIVETPNSEN
GICYPGDFIDYEELREQLSSVSSFERFEIFPKESSWPNHNTNGVTAACSHEGKSSFYRNLL
WLTEKEGSYPKLKNSYVNKKGKEVLVLWGIHHPPNSKEQQNLYQNENAYVSVVTSNYNRRF
TPEIAERPKVRDQAGRMNYYWTLLKPGDTIIFEANGNLIAPMYAFALSRGFGSGIITS

SEQ ID NO: 91 STF2.HA1-2*mut*(PR8)

MAQVINTNSLSLLTQNNLNKSQSALGTAIERLSSGLRINSAKDDAAGQAIANRFTANIKGL
TQASRNANDGISIAQTTEGALNEINNNLQRVRELAVQSANSTNSQSDLDSIQAEITQRLNE
IDRVSGQTQFNGVKVLAQDNTLTIQVGANDGETIDIDLKQINSQTLGLDSLNVQKAYDVKD
TAVTTKAYANNGTTLDVSGLDDAAIKAATGGTNGTASVTGGAVKFDADNNKYFVTIGGFTG
ADAAKNGDYEVNVATDGTVTLAAGATKTTMPAGATTKTEVQELKDTPAVVSADAKNALIAG
GVDATDANGAELVKMSYTDKNGKTIEGGYALKAGDKYYAADYDEATGAIKAKTTSYTAADG
TTKTAANQLGGVDGKTEVVTIDGKTYNASKAAGHDFKAQPELAEAAAKTTENPLQKIDAAL
AQVDALRSDLGAVQNRFNSAITNLGNTVNNLSEARSRIEDSDYATEVSNMSRAQILQQAGT
SVLAQANQVPQNVLSLLAKGAAPLQLGKCNIAGWLLGNPECDPLLPVRSWSDIAETPNSEN
GICYPGDFIDYEELREQLSSVSSFERFEIFPKESSWPNHNTNGVTAACSHEGKSSFYRNLL
WLTEKEGSYPKLKNSYVNKKGKEVLVLWGIHHPPNSKEQQNLYQNENAYVSVVTSNYNRRF
TPEIAERPKVRDQAGRMNYYWTLLKPGDTIIFEANGNLIAPMYAFALSRGFGSGIITS

SEQ ID NO: 92 STF2.HA1-3(PR8)

MAQVINTNSLSLLTQNNLNKSQSALGTAIERLSSGLRINSAKDDAAGQAIANRFTANIKGL
TQASRNANDGISIAQTTEGALNEINNNLQRVRELAVQSANSTNSQSDLDSIQAEITQRLNE
IDRVSGQTQFNGVKVLAQDNTLTIQVGANDGETIDIDLKQINSQTLGLDSLNVQKAYDVKD
TAVTTKAYANNGTTLDVSGLDDAAIKAATGGTNGTASVTGGAVKFDADNNKYFVTIGGFTG
ADAAKNGDYEVNVATDGTVTLAAGATKTTMPAGATTKTEVQELKDTPAVVSADAKNALIAG
GVDATDANGAELVKMSYTDKNGKTIEGGYALKAGDKYYAADYDEATGAIKAKTTSYTAADG
TTKTAANQLGGVDGKTEVVTIDGKTYNASKAAGHDFKAQPELAEAAAKTTENPLQKIDAAL
AQVDALRSDLGAVQNRFNSAITNLGNTVNNLSEARSRIEDSDYATEVSNMSRAQILQQAGT
SVLAQANQVPQNVLSLLANSENGICYPGDFIDYEELREQLSSVSSFERFEIFPKESSWPNH

NTNGVTAACSHEGKSSFYRNLLWLTEKEGSYPKLKNSYVNKKGKEVLVLWGIHHPPNSKEQ
QNLYQNENAYVSVVTSNYNRRFTPEIAERPKVRDQAGRMNYYWTLLKPGDTIIFEANGNLI
APMYAFALSRG

SEQ ID NO: 93 HA1-2his(PR8)

KGIAPLQLGKCNIAGWLLGNPECDPLLPVRSWSYIVETPNSENGICYPGDFIDYEELREQL
SSVSSFERFEIFPKESSWPNHNTNGVTAACSHEGKSSFYRNLLWLTEKEGSYPKLKNSYVN
KKGKEVLVLWGIHHPPNSKEQQNLYQNENAYVSVVTSNYNRRFTPEIAERPKVRDQAGRMN
YYWTLLKPGDTIIFEANGNLIAPMYAFALSRGFGSGIITSHHHHHH

SEQ ID NO: 94 STF2.4xM2e

MAQVINTNSLSLLTQNNLNKSQSALGTAIERLSSGLRINSAKDDAAGQAIANRFTANIKGL
TQASRNANDGISIAQTTEGALNEINNNLQRVRELAVQSANSTNSQSDLDSIQAEITQRLNE
IDRVSGQTQFNGVKVLAQDNTLTIQVGANDGETIDIDLKQINSQTLGLDSLNVQKAYDVKD
TAVTTKAYANNGTTLDVSGLDDAAIKAATGGTNGTASVTGGAVKFDADNNKYFVTIGGFTG
ADAAKNGDYEVNVATDGTVTLAAGATKTTMPAGATTKTEVQELKDTPAVVSADAKNALIAG
GVDATDANGAELVKMSYTDKNGKTIEGGYALKAGDKYYAADYDEATGAIKAKTTSYTAADG
TTKTAANQLGGVDGKTEVVTIDGKTYNASKAAGHDFKAQPELAEAAAKTTENPLQKIDAAL
AQVDALRSDLGAVQNRFNSAITNLGNTVNNLSEARSRIEDSDYATEVSNMSRAQILQQAGT
SVLAQANQVPQNVLSLLRLSLLTEVETPIRNEWGSRSNDSSDPLESLLTEVETPIRNEWGS
RSNDSSDPGSSLLTEVETPIRNEWGSRSNDSSDPELSLLTEVETPIRNEWGSRSNDSSDPS
R

SEQ ID NO: 95 STF2.HA1-2(VN)

MAQVINTNSLSLLTQNNLNKSQSALGTAIERLSSGLRINSAKDDAAGQAIANRFTANIKGL
TQASRNANDGISIAQTTEGALNEINNNLQRVRELAVQSANSTNSQSDLDSIQAEITQRLNE
IDRVSGQTQFNGVKVLAQDNTLTIQVGANDGETIDIDLKQINSQTLGLDSLNVQKAYDVKD
TAVTTKAYANNGTTLDVSGLDDAAIKAATGGTNGTASVTGGAVKFDADNNKYFVTIGGFTG
ADAAKNGDYEVNVATDGTVTLAAGATKTTMPAGATTKTEVQELKDTPAVVSADAKNALIAG
GVDATDANGAELVKMSYTDKNGKTIEGGYALKAGDKYYAADYDEATGAIKAKTTSYTAADG
TTKTAANQLGGVDGKTEVVTIDGKTYNASKAAGHDFKAQPELAEAAAKTTENPLQKIDAAL
AQVDALRSDLGAVQNRFNSAITNLGNTVNNLSEARSRIEDSDYATEVSNMSRAQILQQAGT
SVLAQANQVPQNVLSLLAGVKPLILRDCSVAGWLLGNPMCDEFINVPEWSYIVEKANPVND
LCYPGDFNDYEELKHLLSRINHFEKIQIIPKSSWSSHEASLGVSSACPYQGKSSFFRNVVW
LIKKNSTYPTIKRSYNNTNQEDLLVLWGIHHPNDAAEQTKLYQNPTTYISVGTSTLNQRLV
PRIATRSKVNGQSGRMEFFWTILKPNDAINFESNGNFIAPEYAYKIVKKGDSTIMKSE

SEQ ID NO: 96 STF2

MAQVINTNSLSLLTQNNLNKSQSALGTAIERLSSGLRINSAKDDAAGQAIANRFTANIKGL
TQASRNANDGISIAQTTEGALNEINNNLQRVRELAVQSANSTNSQSDLDSIQAEITQRLNE
IDRVSGQTQFNGVKVLAQDNTLTIQVGANDGETIDIDLKQINSQTLGLDSLNVQKAYDVKD
TAVTTKAYANNGTTLDVSGLDDAAIKAATGGTNGTASVTGGAVKFDADNNKYFVTIGGFTG
ADAAKNGDYEVNVATDGTVTLAAGATKTTMPAGATTKTEVQELKDTPAVVSADAKNALIAG
GVDATDANGAELVKMSYTDKNGKTIEGGYALKAGDKYYAADYDEATGAIKAKTTSYTAADG
TTKTAANQLGGVDGKTEVVTIDGKTYNASKAAGHDFKAQPELAEAAAKTTENPLQKIDAAL
AQVDALRSDLGAVQNRFNSAITNLGNTVNNLSEARSRIEDSDYATEVSNMSRAQILQQAGT
SVLAQANQVPQNVLSLLA

SEQ ID NO: 97 STF2.Blp.wt

ATGAAATTCTTAGTCAACGTTGCCCTTGTTTTTATGGTCGTATACATTTCTTACATCTAT
GCGGCCCAGGTTATCAATACCAACTCCCTGTCGTTGCTCACCCAAAATAACCTTAATAAA
AGCCAGAGCGCACTGGGAACCGCCATAGAACGCCTCTCAAGCGGCCTCCGGATCAATTCT
GCAAAAGACGACGCCGCCGGTCAGGCCATCGCAAACCGCTTTACCGCCAATATCAAGGGA

CTGACGCAGGCTTCGAGGAATGCTAACGATGGAATAAGCATCGCTCAAACCACGGAGGGC
GCCCTGAACGAGATCAACAACAACCTACAGCGCGTCAGGGAGCTCGCAGTGCAGTCCGCC
AATTCGACCAACTCGCAGTCGGACCTGGACTCGATCCAAGCCGAAATCACCCAGCGCCTG
AATGAGATTGACCGGGTGAGCGGTCAGACACAGTTTAACGGCGTGAAGGTACTTGCACAG
GATAACACACTTACGATACAGGTGGGCGCCAACGATGGTGAAACCATAGACATTGATCTC
AAACAGATTAACAGCCAGACGCTCGGGTTGGATAGCCTGAATGTGCAAAAGGCGTACGAC
GTGAAAGACACGGCGGTCACTACCAAAGCCTACGCTAACAATGGCACTACCTTGGATGTG
AGCGGATTGGATGATGCAGCAATCAAGGCTGCTACCGGCGGTACGAACGGAACCGCGTCC
GTGACCGGCGGTGCCGTGAAGTTCGATGCTGACAACAATAAGTATTTCGTCACCATTGGA
GGCTTTACTGGCGCCGACGCAGCAAAGAACGGCGACTATGAAGTGAACGTGGCAACCGAT
GGAACCGTGACGCTGGCCGCTGGTGCCACCAAGACCACCATGCCAGCCGGCGCCACAACT
AAGACCGAGGTGCAGGAGTTAAAGGACACCCCCGCGGTGGTTAGCGCAGATGCCAAAAAC
GCGTTGATCGCCGGCGGAGTGGATGCAACTGATGCTAATGGTGCGGAGCTGGTTAAAATG
TCGTATACAGACAAGAATGGTAAGACGATCGAGGGCGGTTATGCCCTTAAGGCAGGAGAT
AAGTATTACGCTGCTGATTACGATGAGGCGACGGGAGCTATTAAGGCCAAGACAACGTCA
TACACGGCGGCGGACGGAACGACTAAGACGGCTGCCAATCAGTTGGGAGGGGTTGACGGG
AAGACAGAGGTCGTTACGATCGATGGCAAGACATACAACGCCTCCAAGGCCGCTGGCCAC
GATTTCAAAGCTCAACCCGAACTGGCCGAGGCCGCGGCGAAAACAACTGAGAACCCGTTG
CAGAAGATTGATGCGGCCCTGGCGCAAGTAGATGCCCTGCGCTCAGACCTGGGCGCCGTT
CAAAATCGATTCAATTCCGCGATTACAAACCTGGGCAATACAGTAAACAATCTATCCGAG
GCCAGATCCCGCATTGAAGACTCCGACTACGCGACAGAAGTAAGTAACATGAGTCGTGCC
CAGATTCTGCAGCAGGCCGGCACTAGTGTCCTGGCCCAGGCCAATCAAGTCCCGCAGAAT
GTGCTGAGCCTACTACGAGAGTTCAGTCGTTACCCAGCCCAATGGCGGCCGCTC

SEQ ID NO: 98 STF2.Blp.ng

ATGAAATTCTTAGTCAACGTTGCCCTTGTTTTTATGGTCGTATACATTTCTTACATCTAT
GCGGCCCAAGTGATCAACACCAACTCCCTGTCCCTGCTCACTCAAAACAACCTCCAGAAA
TCTCAGTCTGCTCTCGGTACTGCTATCGAGCGTCTCTCCTCCGGCCTGAGGATCAACTCC
GCTAAAGATGACGCCGCCGGACAAGCTATCGCTAACCGCTTTACTGCCAACATTAAAGGC
CTGACCCAAGCCTCTAGGAACGCCAATGATGGCATTTCTATCGCTCAAACCACCGAAGGC
GCCCTGAACGAAATCAATAATAACCTGCAACGTGTCCGCGAACTCGCCGTCCAGTCCGCC
CAATCTACCAACTCTCAGTCTGACCTGGACTCTATCCAAGCTGAGATCACCCAAAGGCTC
AACGAGATCGATCGTGTGTCTGGTCAAACTCAGTTTAACGGCGTCAAAGTGCTGGCTCAA
GACAACACACTCACCATCCAAGTAGGAGCCAATGACGGCGAGACAATCGATATCGACCTG
AAGCAGATCAATTCTCAAACTCTGGGCCTGGACTCCCTGAACGTCCAAAAGGCTTACGAC
GTGAAGGACACCGCTGTGACAACCAAAGCCTATGCTAATCAAGGAACAACCCTGGACGTG
TCCGGACTCGATGACGCCGCTATCAAGGCCGCTACTGGCGGCACTCAGGGAACTGCCTCT
GTGACCGGAGGCGCTGTGAAGTTCGACGCCGATAACAACAAATACTTCGTCACTATTGGT
GGTTTCACTGGCGCTGACGCCGCTAAGAACGGTGACTACGAGGTCAACGTCGCCACTGAC
GGAACAGTGACACTGGCCGCTGGCGCTACCAAGACCACAATGCCTGCTGGTGCTACTACT
AAGACAGAGGTGCAGGAACTCAAGGACACCCCTGCCGTGGTGTCCGCCGATGCTAAGAAC
GCTCTCATTGCTGGTGGTGTCGACGCTACGGACGCCAACGGAGCCGAGCTTGTGAAAATG
TCCTACACCGACAAGAACGGAAAGACTATCGAGGGAGGTTACGCTCTGAAGGCTGGCGAT
AAGTACTACGCCGCTGATTATGACGAAGCTACAGGCGCCATTAAAGCTAAAACCACATCT
TATACCGCTGCCGATGGTACTACCAAGACTGCTGCCAATCAGCTGGGTGGAGTCGATGGA
AAAACCGAAGTGGTCACAATCGACGGAAAGACCTATCAAGCCTCCAAGGCTGCTGGCCAC
GACTTCAAGGCCCAACCCGAACTGGCCGAGGCTGCTGCTAAAACCACTGAAAACCCCCTG
CAGAAAATTGACGCTGCCCTGGCCCAAGTGGATGCCCTCCGCTCCGACCTCGGCGCTGTG
CAAAACCGCTTCAACTCCGCTATTACAAACCTCGGCAATACCGTCAATCAGCTCTCTGAA
GCTAGGTCTCGTATTGAGGATTCCGATTACGCTACCGAGGTCTCCCAGATGTCCCGTGCC
CAAATTCTCCAACAAGCCGGCACCTCCGTCCTCGCCCAAGCCAATCAGGTGCCACAGAAT
GTGCTGAGC

SEQ ID NO: 99 Honey Bee mellitin

ATGAAATTCTTAGTCAACGTTGCCCTTGTTTTTATGGTCGTATACATTTCTTACATCTAT
GCG

SEQ ID NO: 100 STF2.HA0s

ATGAAATTCTTAGTCAACGTTGCCCTTGTTTTTATGGTCGTATACATTTCTTACATCTAT
GCGGCCCAGGTTATCAATACCAACTCCCTGTCGTTGCTCACCCAAAATAACCTTAATAAA
AGCCAGAGCGCACTGGGAACCGCCATAGAACGCCTCTCAAGCGGCCTCCGGATCAATTCT
GCAAAAGACGACGCCGCCGGTCAGGCCATCGCAAACCGCTTTACCGCCAATATCAAGGGA
CTGACGCAGGCTTCGAGGAATGCTAACGATGGAATAAGCATCGCTCAAACCACGGAGGGC
GCCCTGAACGAGATCAACAACAACCTACAGCGCGTCAGGGAGCTCGCAGTGCAGTCCGCC
AATTCGACCAACTCGCAGTCGGACCTGGACTCGATCCAAGCCGAAATCACCCAGCGCCTG
AATGAGATTGACCGGGTGAGCGGTCAGACACAGTTTAACGGCGTGAAGGTACTTGCACAG
GATAACACACTTACGATACAGGTGGGCGCCAACGATGGTGAAACCATAGACATTGATCTC
AAACAGATTAACAGCCAGACGCTCGGGTTGGATAGCCTGAATGTGCAAAAGGCGTACGAC
GTGAAAGACACGGCGGTCACTACCAAAGCCTACGCTAACAATGGCACTACCTTGGATGTG
AGCGGATTGGATGATGCAGCAATCAAGGCTGCTACCGGCGGTACGAACGGAACCGCGTCC
GTGACCGGCGGTGCCGTGAAGTTCGATGCTGACAACAATAAGTATTTCGTCACCATTGGA
GGCTTTACTGGCGCCGACGCAGCAAAGAACGGCGACTATGAAGTGAACGTGGCAACCGAT
GGAACCGTGACGCTGGCCGCTGGTGCCACCAAGACCACCATGCCAGCCGGCGCCACAACT
AAGACCGAGGTGCAGGAGTTAAAGGACACCCCCGCGGTGGTTAGCGCAGATGCCAAAAAC
GCCTTGATCGCCGGCGGAGTGGATGCAACTGATGCTAATGGTGCGGAGCTGGTTAAAATG
TCGTATACAGACAAGAATGGTAAGACGATCGAGGGCGGTTATGCCCTTAAGGCAGGAGAT
AAGTATTACGCTGCTGATTACGATGAGGCGACGGGAGCTATTAAGGCCAAGACAACGTCA
TACACGGCGGCGGACGGAACGACTAAGACGGCTGCCAATCAGTTGGGAGGGGTTGACGGG
AAGACAGAGGTCGTTACGATCGATGGCAAGACATACAACGCTTCCAAGGCCGCTGGCCAC
GATTTCAAAGCTCAACCCGAACTGGCCGAGGCCGCGGCGAAAACAACTGAGAACCCGTTG
CAGAAGATTGATGCGGCCCTGGCGCAAGTAGATGCCCTGCGCTCAGACCTGGGCGCCGTT
CAAAATCGATTCAATTCCGCGATTACAAACCTGGGCAATACAGTAAACAATCTATCCGAG
GCCAGATCCCGCATTGAAGACTCCGACTACGCGACAGAAGTAAGTAACATGAGTCGTGCC
CAGATTCTGCAGCAGGCCGGCACTAGTGTCCTGGCCCAGGCCAATCAAGTCCCGCAGAAT
GTGCTGAGCCTACTAGCATCAGGTTCCGGCTCAGGTTCCGACACCATTTGCATTGGATAC
CATGCAAACAACTCAACCGATACTGTTGATACCGTCCTTGAGAAGAACGTTACCGTCACG
CACTCGGTCAACCTATTAGAGGATAGCCACAACGGAAAGCTGTGCCGTCTGAAAGGCATA
GCGCCACTGCAGCTCGGCAAATGTAACATCGCAGGTTGGCTCCTTGGTAACCCGGAGTGC
GACCCCCTCCTCCCTGTACGATCCTGGAGTTATATCGTGGAGACCCCCAATAGCGAGAAC
GGAATTTGCTACCCAGGAGATTTCATAGACTACGAGGAGTTGCGCGAGCAGCTTTCGTCT
GTGAGCAGCTTCGAAAGGTTCGAGATATTCCCGAAGGAGAGCAGCTGGCCGAATCATAAC
ACTAACGGTGTGACAGCCGCCTGCAGTCATGAAGGAAAGAGTTCATTCTATCGCAACCTG
CTGTGGTTGACGGAGAAAGAGGGCAGCTACCCTAAGTTGAAGAACTCCTATGTGAACAAA
AAAGGCAAGGAGGTTCTGGTGCTGTGGGGCATACACCACCCCCCCAATAGCAAGGAGCAG
CAGAATCTGTACCAAAACGAGAATGCCTATGTGAGCGTGGTCACTAGTAACTATAACCGT
CGGTTCACTCCCGAGATCGCCGAGCGTCCGAAGGTGAGGGACCAGGCAGGCCGGATGAAC
TACTACTGGACCCTATTGAAGCCAGGGGACACGATTATCTTCGAGGCAAACGGAAACCTC
ATAGCGCCGATGTACGCCTTCGCCCTGAGCCGTGGCTTTGGATCGGGGATCATCACGTCT
AACGCCTCGATGCACGAATGTAATACCAAATGCCAGACCCCACTGGGTGCTATCAACTCG
TCCTTACCCTATCAAAATATACATCCGGTCACCATAGGCGAGTGTCCCAAATATGTCAGA
TCCGCCAAGTTGCGGATGGTGACCGGCCTCCGCAATATTCCTAGTATTCAGTCACGCGGC
TTGTTCGGCGCCATCGCTGGTTTCATCGAAGGCGGGTGGACAGGCATGATTGATGGCTGG
TATGGCTATCACCACCAGAACGAGCAGGGCTCGGGCTACGCGGCTGACCAGAAGTCGACT
CAGAATGCCATCAATGGCATCACGAATAAGGTGAACACGGTCATTGAAAAGATGAACATT
CAATTTACAGCCGTAGGAAAAGAGTTTAATAAACTGGAAAAAAGAATGGAGAATCTGAAT
AAGAAGGTGGACGACGGATTTTTGGACATCTGGACGTACAACGCCGAGCTGCTGGTTCTG

CTGGAAAATGAGCGAACACTGGATTTTCATGATTCTAACGTAAAGAATCTGTACGAGAAG
GTGAAGTCCCAACTAAAGAATAATGCCAAGGAAATCGGAAATGGATGCTTTGAGTTTTAC
CACAAGTGCGATAATGAGTGCATGGAATCCGTGCGAAATGGTACATACGATTACCCAAAG
TACTCCGAAGAATCCAAGCTAAATCGCGAAAAGGTTGATGGTGTTAAACTTGAATCCATG
GGTATTTACCAACACCATCATCACCACCATTAATAG

SEQ ID NO: 101 FOR PRIMER
AGTCGCTGAGCCTACTAGCATCAGGTTCCGGCTCAGGTTCCGACACCATTTGCATTGGAT
ACCATGC

SEQ ID NO: 102 REV PRIMER
AGTCGCATGCCTATTAATGGTGGTGATGATGGTGTTGGTAAATACCCATGGATTC

SEQ ID NO: 103 HA0s PR8
GACACCATTTGCATTGGATACCATGCAAACAACTCAACCGATACTGTTGATACCGTCCTT
GAGAAGAACGTTACCGTCACGCACTCGGTCAACCTATTAGAGGATAGCCACAACGGAAAG
CTGTGCCGTCTGAAAGGCATAGCGCCACTGCAGCTCGGCAAATGTAACATCGCAGGTTGG
CTCCTTGGTAACCCGGAGTGCGACCCCCTCCTCCCTGTACGATCCTGGAGTTATATCGTG
GAGACCCCCAATAGCGAGAACGGAATTTGCTACCCAGGAGATTTCATAGACTACGAGGAG
TTGCGCGAGCAGCTTTCGTCTGTGAGCAGCTTCGAAAGGTTCGAGATATTCCCGAAGGAG
AGCAGCTGGCCGAATCATAACACTAACGGTGTGACAGCCGCCTGCAGTCATGAAGGAAAG
AGTTCATTCTATCGCAACCTGCTGTGGTTGACGGAGAAAGAGGGCAGCTACCCTAAGTTG
AAGAACTCCTATGTGAACAAAAAAGGCAAGGAGGTTCTGGTGCTGTGGGGCATACACCAC
CCCCCCAATAGCAAGGAGCAGCAGAATCTGTACCAAAACGAGAATGCCTATGTGAGCGTG
GTCACTAGTAACTATAACCGTCGGTTCACTCCCGAGATCGCCGAGCGTCCGAAGGTGAGG
GACCAGGCAGGCCGGATGAACTACTACTGGACCCTATTGAAGCCAGGGGACACGATTATC
TTCGAGGCAAACGGAAACCTCATAGCGCCGATGTACGCCTTCGCCCTGAGCCGTGGCTTT
GGATCGGGGATCATCACGTCTAACGCCTCGATGCACGAATGTAATACCAAATGCCAGACC
CCACTGGGTGCTATCAACTCGTCCTTACCCTATCAAAATATACATCCGGTCACCATAGGC
GAGTGTCCCAAATATGTCAGATCCGCCAAGTTGCGGATGGTGACCGGCCTCCGCAATATT
CCTAGTATTCAGTCACGCGGCTTGTTCGGCGCCATCGCTGGTTTCATCGAAGGCGGGTGG
ACAGGCATGATTGATGGCTGGTATGGCTATCACCACCAGAACGAGCAGGGCTCGGGCTAC
GCGGCTGACCAGAAGTCGACTCAGAATGCCATCAATGGCATCACGAATAAGGTGAACACG
GTCATTGAAAAGATGAACATTCAATTTACAGCCGTAGGAAAAGAGTTTAATAAACTGGAA
AAAAGAATGGAGAATCTGAATAAGAAGGTGGACGACGGATTTTTGGACATCTGGACGTAC
AACGCCGAGCTGCTGGTTCTGCTGGAAAATGAGCGAACACTGGATTTTCATGATTCTAAC
GTAAAGAATCTGTACGAGAAGGTGAAGTCCCAACTAAAGAATAATGCCAAGGAAATCGGA
AATGGATGCTTTGAGTTTTACCACAAGTGCGATAATGAGTGCATGGAATCCGTGCGAAAT
GGTACATACGATTACCCAAAGTACTCCGAAGAATCCAAGCTAAATCGCGAAAAGGTTGAT
GGTGTTAAACTTGAATCCATGGGTATTTACCAACACCATCATCACCACCATTAATAG

SEQ ID NO: 104 STF2.HA1-1
ATGAAATTCTTAGTCAACGTTGCCCTTGTTTTTATGGTCGTATACATTTCTTACATCTAT
GCGGCCCAGGTTATCAATACCAACTCCCTGTCGTTGCTCACCCAAAATAACCTTAATAAA
AGCCAGAGCGCACTGGGAACCGCCATAGAACGCCTCTCAAGCGGCCTCCGGATCAATTCT
GCAAAAGACGACGCCGCCGGTCAGGCCATCGCAAACCGCTTTACCGCCAATATCAAGGGA
CTGACGCAGGCTTCGAGGAATGCTAACGATGGAATAAGCATCGCTCAAACCACGGAGGGC
GCCCTGAACGAGATCAACAACAACCTACAGCGCGTCAGGGAGCTCGCAGTGCAGTCCGCC
AATTCGACCAACTCGCAGTCGGACCTGGACTCGATCCAAGCCGAAATCACCCAGCGCCTG
AATGAGATTGACCGGGTGAGCGGTCAGACACAGTTTAACGGCGTGAAGGTACTTGCACAG
GATAACACACTTACGATACAGGTGGGCGCCAACGATGGTGAAACCATAGACATTGATCTC
AAACAGATTAACAGCCAGACGCTCGGGTTGGATAGCCTGAATGTGCAAAAGGCGTACGAC
GTGAAAGACACGGCGGTCACTACCAAAGCCTACGCTAACAATGGCACTACCTTGGATGTG

AGCGGATTGGATGATGCAGCAATCAAGGCTGCTACCGGCGGTACGAACGGAACCGCGTCC
GTGACCGGCGGTGCCGTGAAGTTCGATGCTGACAACAATAAGTATTTCGTCACCATTGGA
GGCTTTACTGGCGCCGACGCAGCAAAGAACGGCGACTATGAAGTGAACGTGGCAACCGAT
GGAACCGTGACGCTGGCCGCTGGTGCCACCAAGACCACCATGCCAGCCGGCGCCACAACT
AAGACCGAGGTGCAGGAGTTAAAGGACACCCCCGCGGTGGTTAGCGCAGATGCCAAAAAC
GCCTTGATCGCCGGCGGAGTGGATGCAACTGATGCTAATGGTGCGGAGCTGGTTAAAATG
TCGTATACAGACAAGAATGGTAAGACGATCGAGGGCGGTTATGCCCTTAAGGCAGGAGAT
AAGTATTACGCTGCTGATTACGATGAGGCGACGGGAGCTATTAAGGCCAAGACAACGTCA
TACACGGCGGCGGACGGAACGACTAAGACGGCTGCCAATCAGTTGGGAGGGGTTGACGGG
AAGACAGAGGTCGTTACGATCGATGGCAAGACATACAACGCTTCCAAGGCCGCTGGCCAC
GATTTCAAAGCTCAACCCGAACTGGCCGAGGCCGCGGCGAAAACAACTGAGAACCCGTTG
CAGAAGATTGATGCGGCCCTGGCGCAAGTAGATGCCCTGCGCTCAGACCTGGGCGCCGTT
CAAAATCGATTCAATTCCGCGATTACAAACCTGGGCAATACAGTAAACAATCTATCCGAG
GCCAGATCCCGCATTGAAGACTCCGACTACGCGACAGAAGTAAGTAACATGAGTCGTGCC
CAGATTCTGCAGCAGGCCGGCACTAGTGTCCTGGCCCAGGCCAATCAAGTCCCGCAGAAT
GTGCTGAGCCTACTAGCATCAGGTTCCGGCTCAGGTTCCAGCCACAACGGAAAGCTGTGC
CGTCTGAAAGGCATAGCGCCACTGCAGCTCGGCAAATGTAACATCGCAGGTTGGCTCCTT
GGTAACCCGGAGTGCGACCCCCTCCTCCCTGTACGATCCTGGAGTTATATCGTGGAGACC
CCCAATAGCGAGAACGGAATTTGCTACCCAGGAGATTTCATAGACTACGAGGAGTTGCGC
GAGCAGCTTTCGTCTGTGAGCAGCTTCGAAAGGTTCGAGATATTCCCGAAGGAGAGCAGC
TGGCCGAATCATAACACTAACGGTGTGACAGCCGCCTGCAGTCATGAAGGAAAGAGTTCA
TTCTATCGCAACCTGCTGTGGTTGACGGAGAAAGAGGGCAGCTACCCTAAGTTGAAGAAC
TCCTATGTGAACAAAAAAGGCAAGGAGGTTCTGGTGCTGTGGGGCATACACCACCCCCCC
AATAGCAAGGAGCAGCAGAATCTGTACCAAAACGAGAATGCCTATGTGAGCGTGGTCACT
AGTAACTATAACCGTCGGTTCACTCCCGAGATCGCCGAGCGTCCGAAGGTGAGGGACCAG
GCAGGCCGGATGAACTACTACTGGACCCTATTGAAGCCAGGGGACACGATTATCTTCGAG
GCAAACGGAAACCTCATAGCGCCGATGTACGCCTTCGCCCTGAGCCGTGGCTTTGGATCG
GGGATCATCACGTCTAACGCCTCGATGCACGAATGTAATACCAAATGCCAGACCCCACTG
GGTGCTATCAACTCGTCCTTACCCTATCAAAATATACATCCGGTCACCATAGGCGAGTGT
CCCAAATATGTCAGACACCATCATCACCACCATTAATAG

SEQ ID NO: 105 FOR PRIMER
AGTCGCTGAGCCTACTAGCATCAGGTTCCGGCTCAGGTTCCAGCCACAACGGAAAGCTGT
GCCGTC

SEQ ID NO: 106 REV PRIMER
AGTCGCATGCCTATTAATGGTGGTGATGATGGTGTCTGACATATTTGGGACACTCGCCTA
TGGTGAC

SEQ ID NO: 107 STF2.HA1-2
ATGAAATTCTTAGTCAACGTTGCCCTTGTTTTTATGGTCGTATACATTTCTTACATCTAT
GCGGCCCAGGTTATCAATACCAACTCCCTGTCGTTGCTCACCCAAAATAACCTTAATAAA
AGCCAGAGCGCACTGGGAACCGCCATAGAACGCCTCTCAAGCGGCCTCCGGATCAATTCT
GCAAAAGACGACGCCGCCGGTCAGGCCATCGCAAACCGCTTTACCGCCAATATCAAGGGA
CTGACGCAGGCTTCGAGGAATGCTAACGATGGGATAAGCATCGCTCAAACCACGGAGGGC
GCCCTGAACGAGATCAACAACAACCTACAGCGCGTCAGGGAGCTCGCAGTGCAGTCCGCC
AATTCGACCAACTCGCAGTCGGACCTGGACTCGATCCAAGCCGAAATCACCCAGCGCCTG
AACGAGATTGACCGGGTGAGCGGTCAGACACAGTTTAACGGCGTGAAGGTACTTGCGCAG
GATAACACACTTACGATACAGGTGGGCGCCAACGATGGTGAAACCATAGACATTGATCTC
AAACAGATTAACAGCCAGACGCTCGGGTTGGATAGCCTGAATGTGCAAAAGGCGTACGAC
GTGAAAGACACGGCGGTCACTACCAAAGCCTACGCTAACAATGGCACTACCTTGGATGTG
AGCGGATTGGATGATGCAGCAATCAAGGCTGCTACCGGCGGTACGAACGGAACCGCGTCC
GTGACCGGCGGTGCCGTGAAGTTCGATGCTGACAACAATAAGTATTTCGTCACCATTGGA

GGCTTTACTGGCGCCGACGCAGCAAAGAACGGCGACTATGAAGTGAACGTGGCAACCGAT
GGAACCGTGACGCTGGCCGCTGGTGCCACCAAGACCACCATGCCAGCCGGCGCCACAACT
AAGACCGAGGTGCAGGAGTTAAAGGACACCCCCGCGGTGGTTAGCGCAGATGCCAAAAAC
GCGTTGATCGCCGGCGGAGTGGATGCAACTGATGCTAATGGTGCGGAGCTGGTTAAAATG
TCGTATACAGACAAGAATGGTAAGACGATCGAGGGCGGTTATGCCCTTAAGGCAGGAGAT
AAGTATTACGCTGCTGATTACGATGAGGCGACGGGAGCTATTAAGGCCAAGACAACGTCA
TACACGGCGGCGGACGGAACGACTAAGACGGCTGCCAATCAGTTGGGAGGGGTTGACGGG
AAGACAGAGGTCGTTACGATCGATGGCAAGACATACAACGCCTCCAAGGCCGCTGGCCAC
GATTTCAAAGCTCAACCCGAACTGGCCGAGGCCGCGGCGAAAACAACTGAGAACCCGTTG
CAGAAGATTGATGCGGCCCTGGCGCAAGTAGATGCCCTGCGCTCAGACCTGGGCGCCGTT
CAAAATCGATTCAATTCCGCGATTACAAACCTGGGCAATACAGTAAACAATCTATCCGAG
GCCAGATCCCGCATTGAAGACTCCGACTACGCGACAGAAGTAAGTAACATGAGTCGTGCC
CAGATTCTGCAGCAGGCCGGCACTAGTGTCCTGGCCCAGGCCAATCAAGTCCCGCAGAAT
GTGCTGAGCCTACTAGCATCAGGTTCCGGCTCAGGTTCCAAAGGCATAGCGCCACTGCAG
CTCGGCAAATGTAACATCGCAGGTTGGCTCCTTGGTAACCCGGAGTGCGACCCCCTCCTC
CCTGTACGATCCTGGAGTTATATCGTGGAGACCCCCAATAGCGAGAACGGAATTTGCTAC
CCAGGAGATTTCATAGACTACGAGGAGTTGCGCGAGCAGCTTTCGTCTGTGAGCAGCTTC
GAGAGGTTCGAAATCTTCCCGAAGGAGAGCAGCTGGCCGAATCATAACACTAACGGTGTG
ACAGCCGCCTGCAGTCATGAAGGAAAGAGTTCATTCTATCGCAACCTGCTGTGGTTGACG
GAGAAAGAGGGCAGCTACCCTAAGTTGAAGAACTCCTATGTGAACAAAAAAGGCAAGGAG
GTTCTGGTGCTGTGGGGCATACACCACCCCCCCAATAGCAAGGAGCAGCAGAATCTGTAC
CAAAACGAGAATGCCTATGTGAGCGTGGTCACTAGTAACTATAACCGTCGGTTCACTCCC
GAGATCGCCGAGCGTCCGAAGGTGAGGGACCAGGCAGGCCGGATGAACTACTACTGGACC
CTATTGAAGCCAGGGGACACGATTATCTTCGAGGCAAACGGAAACCTCATAGCGCCGATG
TACGCGTTCGCCCTGAGCCGTGGCTTTGGATCGGGGATCATCACGTCTCACCATCATCAC
CACCATTAATAG

SEQ ID NO: 108 FOR PRIMER
AGTCGCTGAGCCTACTAGCATCAGGTTCCGGCTCAGGTTCCAAAGGCATAGCGCCACTGC
AGCTCG

SEQ ID NO: 109 REV PRIMER
AGTCGCATGCCTATTAATGGTGGTGATGATGGTGAGACGTGATGATCCCCGATCCAAAGC
CACGGCTCAG

SEQ ID NO: 110 STF2.HA1-2*mut*
ATGAAATTCTTAGTCAACGTTGCCCTTGTTTTTATGGTCGTATACATTTCTTACATCTAT
GCGGCCCAGGTTATCAATACCAACTCCCTGTCGTTGCTCACCCAAAATAACCTTAATAAA
AGCCAGAGCGCACTGGGAACCGCCATAGAACGCCTCTCAAGCGGCCTCCGGATCAATTCT
GCAAAAGACGACGCCGCCGGTCAGGCCATCGCAAACCGCTTTACCGCCAATATCAAGGGA
CTGACGCAGGCTTCGAGGAATGCTAACGATGGAATAAGCATCGCTCAAACCACGGAGGGC
GCCCTGAACGAGATCAACAACAACCTACAGCGCGTCAGGGAGCTCGCAGTGCAGTCCGCC
AATTCGACCAACTCGCAGTCGGACCTGGACTCGATCCAAGCCGAAATCACCCAGCGCCTG
AATGAGATTGACCGGGTGAGCGGTCAGACACAGTTTAACGGCGTGAAGGTACTTGCACAG
GATAACACACTTACGATACAGGTGGGCGCCAACGATGGTGAAACCATAGACATTGATCTC
AAACAGATTAACAGCCAGACGCTCGGGTTGGATAGCCTGAATGTGCAAAAGGCGTACGAC
GTGAAAGACACGGCGGTCACTACCAAAGCCTACGCTAACAATGGCACTACCTTGGATGTG
AGCGGATTGGATGATGCAGCAATCAAGGCTGCTACCGGCGGTACGAACGGAACCGCGTCC
GTGACCGGCGGTGCCGTGAAGTTCGATGCTGACAACAATAAGTATTTCGTCACCATTGGA
GGCTTTACTGGCGCCGACGCAGCAAAGAACGGCGACTATGAAGTGAACGTGGCAACCGAT
GGAACCGTGACGCTGGCCGCTGGTGCCACCAAGACCACCATGCCAGCCGGCGCCACAACT
AAGACCGAGGTGCAGGAGTTAAAGGACACCCCCGCGGTGGTTAGCGCAGATGCCAAAAAC
GCGTTGATCGCCGGCGGAGTGGATGCAACTGATGCTAATGGTGCGGAGCTGGTTAAAATG

TCGTATACAGACAAGAATGGTAAGACGATCGAGGGCGGTTATGCCCTTAAGGCAGGAGAT
AAGTATTACGCTGCTGATTACGATGAGGCGACGGGAGCTATTAAGGCCAAGACAACGTCA
TACACGGCGGCGGACGGAACGACTAAGACGGCTGCCAATCAGTTGGGAGGGGTTGACGGG
AAGACAGAGGTCGTTACGATCGATGGCAAGACATACAACGCCTCCAAGGCCGCTGGCCAC
GATTTCAAAGCTCAACCCGAACTGGCCGAGGCCGCGGCGAAAACAACTGAGAACCCGTTG
CAGAAGATTGATGCGGCCCTGGCGCAAGTAGATGCCCTGCGCTCAGACCTGGGCGCCGTT
CAAAATCGATTCAATTCCGCGATTACAAACCTGGGCAATACAGTAAACAATCTATCCGAG
GCCAGATCCCGCATTGAAGACTCCGACTACGCGACAGAAGTAAGTAACATGAGTCGTGCC
CAGATTCTGCAGCAGGCCGGCACTAGTGTCCTGGCCCAGGCCAATCAAGTCCCGCAGAAT
GTGCTGAGCCTGCTGGCTTCTGGATCTGGATCTGGTTCTAAAGGCGCTGCCCCTCTGCAA
CTCGGCAAGTGCAATATTGCGGGGTGGCTGTTGGGCAACCCAGAATGTGACCCCCTCCTG
CCCGTCCGTTCTTGGTCTGACATCGCTGAAACACCTAACTCCGAGAACGGCATCTGTTAC
CCGGGCGACTTCATTGACTACGAGGAGCTCCGCGAACAGCTGTCTTCTGTCTCTTCTTTC
GAACGTTTCGAAATTTTCCCAAAGGAATCCTCCTGGCCAAACCATAACACCAACGGAGTG
ACCGCCGCCTGTAGCCATGAGGGCAAGTCTTCTTTCTACCGTAATCTGCTGTGGCTGACT
GAAAAAGAGGGTTCTTATCCCAAACTGAAGAACTCTTATGTCAACAAGAAGGGCAAAGAG
GTCCTGGTGCTGTGGGGTATCCACCACCCCCCCAACTCCAAGGAGCAGCAGAATCTGTAC
CAAAACGAAAATGCTTACGTGTCTGTCGTGACATCTAACTACAACCGGCGCTTCACGCCC
GAAATCGCCGAGCGTCCCAAAGTGCGCGATCAGGCCGGAAGGATGAACTACTACTGGACC
CTGCTCAAACCCGGAGATACCATCATATTCGAAGCCAATGGCAATCTCATCGCTCCCATG
TATGCTTTCGCTCTCTCTCGCGGATTCGGTTCCGGCATAATCACTAGCCACCACCACCAC
CACCACTAATAG

SEQ ID NO: 111 HA1-2*mut*

AAGGCGCTGCCCCTCTGCAACTCGGCAAGTGCAATATTGCGGGGTGGCTGTTGGGCAACC
CAGAATGTGACCCCCTCCTGCCCGTCCGTTCTTGGTCTGACATCGCTGAAACACCTAACT
CCGAGAACGGCATCTGTTACCCGGGCGACTTCATTGACTACGAGGAGCTCCGCGAACAGC
TGTCTTCTGTCTCTTCTTTCGAACGTTTCGAAATTTTCCCAAAGGAATCCTCCTGGCCAA
ACCATAACACCAACGGAGTGACCGCCGCCTGTAGCCATGAGGGCAAGTCTTCTTTCTACC
GTAATCTGCTGTGGCTGACTGAAAAAGAGGGTTCTTATCCCAAACTGAAGAACTCTTATG
TCAACAAGAAGGGCAAAGAGGTCCTGGTGCTGTGGGGTATCCACCACCCCCCCAACTCCA
AGGAGCAGCAGAATCTGTACCAAAACGAAAATGCTTACGTGTCTGTCGTGACATCTAACT
ACAACCGGCGCTTCACGCCCGAAATCGCCGAGCGTCCCAAAGTGCGCGATCAGGCCGGAA
GGATGAACTACTACTGGACCCTGCTCAAACCCGGAGATACCATCATATTCGAAGCCAATG
GCAATCTCATCGCTCCCATGTATGCTTTCGCTCTCTCTCGCGGATTCGGTTCCGGCATAA
TCACTAGCCACCACCACCACCACCACTAATAG

SEQ ID NO: 112 STF2.HA1-3

ATGAAATTCTTAGTCAACGTTGCCCTTGTTTTTATGGTCGTATACATTTCTTACATCTAT
GCGGCCCAGGTTATCAATACCAACTCCCTGTCGTTGCTCACCCAAAATAACCTTAATAAA
AGCCAGAGCGCACTGGGAACCGCCATAGAACGCCTCTCAAGCGGCCTCCGGATCAATTCT
GCAAAAGACGACGCCGCCGGTCAGGCCATCGCAAACCGCTTTACCGCCAATATCAAGGGA
CTGACGCAGGCTTCGAGGAATGCTAACGATGGGATAAGCATCGCTCAAACCACGGAGGGC
GCCCTGAACGAGATCAACAACAACCTACAGCGCGTCAGGGAGCTCGCAGTGCAGTCCGCC
AATTCGACCAACTCGCAGTCGGACCTGGACTCGATCCAAGCCGAAATCACCCAGCGCCTG
AACGAGATTGACCGGGTGAGCGGTCAGACACAGTTTAACGGCGTGAAGGTACTTGCGCAG
GATAACACACTTACGATACAGGTGGGCGCCAACGATGGTGAAACCATAGACATTGATCTC
AAACAGATTAACAGCCAGACGCTCGGGTTGGATAGCCTGAATGTGCAAAAGGCGTACGAC
GTGAAAGACACGGCGGTCACTACCAAAGCCTACGCTAACAATGGCACTACCTTGGATGTG
AGCGGATTGGATGATGCAGCAATCAAGGCTGCTACCGGCGGTACGAACGGAACCGCGTCC
GTGACCGGCGGTGCCGTGAAGTTCGATGCTGACAACAATAAGTATTTCGTCACCATTGGA
GGCTTTACTGGCGCCGACGCAGCAAAGAACGGCGACTATGAAGTGAACGTGGCAACCGAT
GGAACCGTGACGCTGGCCGCTGGTGCCACCAAGACCACCATGCCAGCCGGCGCCACAACT

AAGACCGAGGTGCAGGAGTTAAAGGACACCCCCGCGGTGGTTAGCGCAGATGCCAAAAAC
GCGTTGATCGCCGGCGGAGTGGATGCAACTGATGCTAATGGTGCGGAGCTGGTTAAAATG
TCGTATACAGACAAGAATGGTAAGACGATCGAGGGCGGTTATGCCCTTAAGGCAGGAGAT
AAGTATTACGCTGCTGATTACGATGAGGCGACGGGAGCTATTAAGGCCAAGACAACGTCA
TACACGGCGGCGGACGGAACGACTAAGACGGCTGCCAATCAGTTGGGAGGGGTTGACGGG
AAGACAGAGGTCGTTACGATCGATGGCAAGACATACAACGCCTCCAAGGCCGCTGGCCAC
GATTTCAAAGCTCAACCCGAACTGGCCGAGGCCGCGGCGAAAACAACTGAGAACCCGTTG
CAGAAGATTGATGCGGCCCTGGCGCAAGTAGATGCCCTGCGCTCAGACCTGGGCGCCGTT
CAAAATCGATTCAATTCCGCGATTACAAACCTGGGCAATACAGTAAACAATCTATCCGAG
GCCAGATCCCGCATTGAAGACTCCGACTACGCGACAGAAGTAAGTAACATGAGTCGTGCC
CAGATTCTGCAGCAGGCCGGCACTAGTGTCCTGGCCCAGGCCAATCAAGTCCCGCAGAAT
GTGCTGAGCCTACTAGCATCAGGTTCCGGCTCAGGTTCCAATAGCGAGAACGGAATTTGC
TACCCAGGAGATTTCATAGACTACGAGGAGTTGCGCGAGCAGCTTTCGTCTGTGAGCAGC
TTCGAGAGGTTCGAAATCTTCCCGAAGGAGAGCAGCTGGCCGAATCATAACACTAACGGT
GTGACAGCCGCCTGCAGTCATGAAGGAAAGAGTTCATTCTATCGCAACCTGCTGTGGTTG
ACGGAGAAGGAGGGCAGCTACCCTAAGCTGAAGAACTCCTATGTGAACAAAAAAGGCAAG
GAGGTTCTGGTGCTGTGGGGCATACACCACCCCCCCAATAGCAAGGAGCAGCAGAATCTG
TACCAAAACGAGAATGCCTATGTGAGCGTGGTCACTAGTAACTATAACCGTCGGTTCACT
CCCGAGATCGCCGAGCGTCCGAAGGTGAGGGACCAGGCAGGCCGGATGAACTACTACTGG
ACCCTATTGAAGCCAGGGGACACGATTATCTTCGAGGCAAACGGAAACCTCATAGCGCCG
ATGTACGCCTTCGCCCTGAGCCGTGGCCACCATCATCACCACCATTAATAG

SEQ ID NO: 113 FOR PRIMER
AGTCGCTGAGCCTACTAGCATCAGGTTCCGGCTCAGGTTCCAATAGCGAGAACGGAATTT
GCTACC

SEQ ID NO: 114 REV PRIMER
AGTCGCATGCCTATTAATGGTGGTGATGATGGTGGCCACGGCTCAGGGCGAAGG

SEQ ID NO: 115 STF2.HA1-3*mut*
ATGAAATTCTTAGTCAACGTTGCCCTTGTTTTTATGGTCGTATACATTTCTTACATCTAT
GCGGCCCAGGTTATCAATACCAACTCCCTGTCGTTGCTCACCCAAAATAACCTTAATAAA
AGCCAGAGCGCACTGGGAACCGCCATAGAACGCCTCTCAAGCGGCCTCCGGATCAATTCT
GCAAAAGACGACGCCGCCGGTCAGGCCATCGCAAACCGCTTTACCGCCAATATCAAGGGA
CTGACGCAGGCTTCGAGGAATGCTAACGATGGAATAAGCATCGCTCAAACCACGGAGGGC
GCCCTGAACGAGATCAACAACAACCTACAGCGCGTCAGGGAGCTCGCAGTGCAGTCCGCC
AATTCGACCAACTCGCAGTCGGACCTGGACTCGATCCAAGCCGAAATCACCCAGCGCCTG
AATGAGATTGACCGGGTGAGCGGTCAGACACAGTTTAACGGCGTGAAGGTACTTGCACAG
GATAACACACTTACGATACAGGTGGGCGCCAACGATGGTGAAACCATAGACATTGATCTC
AAACAGATTAACAGCCAGACGCTCGGGTTGGATAGCCTGAATGTGCAAAAGGCGTACGAC
GTGAAAGACACGGCGGTCACTACCAAAGCCTACGCTAACAATGGCACTACCTTGGATGTG
AGCGGATTGGATGATGCAGCAATCAAGGCTGCTACCGGCGGTACGAACGGAACCGCGTCC
GTGACCGGCGGTGCCGTGAAGTTCGATGCTGACAACAATAAGTATTTCGTCACCATTGGA
GGCTTTACTGGCGCCGACGCAGCAAAGAACGGCGACTATGAAGTGAACGTGGCAACCGAT
GGAACCGTGACGCTGGCCGCTGGTGCCACCAAGACCACCATGCCAGCCGGCGCCACAACT
AAGACCGAGGTGCAGGAGTTAAAGGACACCCCCGCGGTGGTTAGCGCAGATGCCAAAAAC
GCGTTGATCGCCGGCGGAGTGGATGCAACTGATGCTAATGGTGCGGAGCTGGTTAAAATG
TCGTATACAGACAAGAATGGTAAGACGATCGAGGGCGGTTATGCCCTTAAGGCAGGAGAT
AAGTATTACGCTGCTGATTACGATGAGGCGACGGGAGCTATTAAGGCCAAGACAACGTCA
TACACGGCGGCGGACGGAACGACTAAGACGGCTGCCAATCAGTTGGGAGGGGTTGACGGG
AAGACAGAGGTCGTTACGATCGATGGCAAGACATACAACGCCTCCAAGGCCGCTGGCCAC
GATTTCAAAGCTCAACCCGAACTGGCCGAGGCCGCGGCGAAAACAACTGAGAACCCGTTG
CAGAAGATTGATGCGGCCCTGGCGCAAGTAGATGCCCTGCGCTCAGACCTGGGCGCCGTT

CAAAATCGATTCAATTCCGCGATTACAAACCTGGGCAATACAGTAAACAATCTATCCGAG
GCCAGATCCCGCATTGAAGACTCCGACTACGCGACAGAAGTAAGTAACATGAGTCGTGCC
CAGATTCTGCAGCAGGCCGGCACTAGTGTCCTGGCCCAGGCCAATCAAGTCCCGCAGAAT
GTGCTGAGCCTGCTGGCTTCTGGATCTGGATCTGGTTCTAACTCCGAGAACGAAATCTGT
TACCCGGGCGACTTCATTGACAAAGAGGAGCTCCGCGAACAGCTGTCTTCTGTCTCTTCT
TTCGAACGTTTCGAAATTTTCCCAAAGGAATCCTCCTGGCCAAACCATAACACCAACGGA
GTGACCGCCGCCTGTAGCCATGAGGGCAAGTCTTCTTTCTACCGTAATCTGCTGTGGCTG
ACTGAAAAAGAGGGTTCTTATCCCAAACTGAAGAACTCTTATGTCAACAAGAAGGGCAAA
GAGGTCCTGGTGCTGTGGGGTATCCACCACCCCCCCAACTCCAAGGAGCAGCAGAATCTG
TACCAAAACGAAAATGCTTACGTGTCTGTCGTGACATCTAACTACAACCGGCGCTTCACG
CCCGAAATCGCCGAGCGTCCCAAAGTGCGCGATCAGGCCGGAAGGATGAACTACTACTGG
ACCCTGCTCAAACCCGGAGATACCATCATATTCGAAGCCAATGGCAATCTCATCGCTCCC
ATGTATGCTGCTGCTCTCTCTCGCGGACACCACCACCACCACCACTAATAG

SEQ ID NO: 116 HA1-3*mut* PR8

AACTCCGAGAACGAAATCTGTTACCCGGGCGACTTCATTGACAAAGAGGAGCTCCGCGAA
CAGCTGTCTTCTGTCTCTTCTTTCGAACGTTTCGAAATTTTCCCAAAGGAATCCTCCTGG
CCAAACCATAACACCAACGGAGTGACCGCCGCCTGTAGCCATGAGGGCAAGTCTTCTTTC
TACCGTAATCTGCTGTGGCTGACTGAAAAAGAGGGTTCTTATCCCAAACTGAAGAACTCT
TATGTCAACAAGAAGGGCAAAGAGGTCCTGGTGCTGTGGGGTATCCACCACCCCCCCAAC
TCCAAGGAGCAGCAGAATCTGTACCAAAACGAAAATGCTTACGTGTCTGTCGTGACATCT
AACTACAACCGGCGCTTCACGCCCGAAATCGCCGAGCGTCCCAAAGTGCGCGATCAGGCC
GGAAGGATGAACTACTACTGGACCCTGCTCAAACCCGGAGATACCATCATATTCGAAGCC
AATGGCAATCTCATCGCTCCCATGTATGCTGCTGCTCTCTCTCGCGGACACCACCACCAC
CACCACTAATAG

SEQ ID NO: 117 ngSTF2.HA0s

ATGAAATTCTTAGTCAACGTTGCCCTTGTTTTTATGGTCGTATACATTTCTTACATCTAT
GCGGCCCAAGTGATCAACACCAACTCCCTGTCCCTGCTCACTCAAAACAACCTCCAGAAA
TCTCAGTCTGCTCTCGGTACTGCTATCGAGCGTCTCTCCTCCGGCCTGAGGATCAACTCC
GCTAAAGATGACGCCGCCGGACAAGCTATCGCTAACCGCTTTACTGCCAACATTAAAGGC
CTGACCCAAGCCTCTAGGAACGCCAATGATGGCATTTCTATCGCTCAAACCACCGAAGGC
GCCCTGAACGAAATCAATAATAACCTGCAACGTGTCCGCGAACTCGCCGTCCAGTCCGCC
CAATCTACCAACTCTCAGTCTGACCTGGACTCTATCCAAGCTGAGATCACCCAAAGGCTC
AACGAGATCGATCGTGTGTCTGGTCAAACTCAGTTTAACGGCGTCAAAGTGCTGGCTCAA
GACAACACACTCACCATCCAAGTAGGAGCCAATGACGGCGAGACAATCGATATCGACCTG
AAGCAGATCAATTCTCAAACTCTGGGCCTGGACTCCCTGAACGTCCAAAAGGCTTACGAC
GTGAAGGACACCGCTGTGACAACCAAAGCCTATGCTAATCAAGGAACAACCCTGGACGTG
TCCGGACTCGATGACGCCGCTATCAAGGCCGCTACTGGCGGCACTCAGGGAACTGCCTCT
GTGACCGGAGGCGCTGTGAAGTTCGACGCCGATAACAACAAATACTTCGTCACTATTGGT
GGTTTCACTGGCGCTGACGCCGCTAAGAACGGTGACTACGAGGTCAACGTCGCCACTGAC
GGAACAGTGACACTGGCCGCTGGCGCTACCAAGACCACAATGCCTGCTGGTGCTACTACT
AAGACAGAGGTGCAGGAACTCAAGGACACCCCTGCCGTGGTGTCCGCCGATGCTAAGAAC
GCTCTCATTGCTGGTGGTGTCGACGCTACGGACGCCAACGGAGCCGAGCTTGTGAAAATG
TCCTACACCGACAAGAACGGAAAGACTATCGAGGGAGGTTACGCTCTGAAGGCTGGCGAT
AAGTACTACGCCGCTGATTATGACGAAGCTACAGGCGCCATTAAAGCTAAAACCACATCT
TATACCGCTGCCGATGGTACTACCAAGACTGCTGCCAATCAGCTGGGTGGAGTCGATGGA
AAAACCGAAGTGGTCACAATCGACGGAAAGACCTATCAAGCCTCCAAGGCTGCTGGCCAC
GACTTCAAGGCCCAACCCGAACTGGCCGAGGCTGCTGCTAAAACCACTGAAAACCCCCTG
CAGAAAATTGACGCTGCCCTGGCCCAAGTGGATGCCCTCCGCTCCGACCTCGGCGCTGTG
CAAAACCGCTTCAACTCCGCTATTACAAACCTCGGCAATACCGTCAATCAACTCTCTGAA
GCTAGGTCTCGTATTGAGGATTCCGATTACGCTACCGAGGTCTCCCAGATGTCCCGTGCC
CAAATTCTCCAACAAGCCGGCACCTCCGTCCTCGCCCAAGCCAATCAGGTGCCACAGAAT

GTGCTGAGCCTGCTGGCTTCTGGATCTGGATCTGGTTCTGACACCATCTGCATTGGCTAC
CATGCCCAGCAGTCTACAGATACAGTCGATACCGTCCTGGAGAAACAAGTCACCGTCACG
CACTCCGTGAACCTCCTGGAGGACTCCCATAATGGCAAATTGTGTCGGCTGAAAGGCATC
GCCCCTCTGCAACTCGGCAAGTGCAATATTGCGGGGTGGCTGTTGGGCAACCCAGAATGT
GACCCCCTCCTGCCCGTCCGTTCTTGGTCTTACATCGTGGAAACACCTAACTCCGAGAAC
GGCATCTGTTACCCGGGCGACTTCATTGACTACGAGGAGCTCCGCGAACAGCTGTCTTCT
GTCTCTTCTTTCGAACGTTTCGAAATTTTCCCAAAGGAATCCTCCTGGCCAAACCATAAC
ACCAACGGAGTGACCGCCGCCTGTAGCCATGAGGGCAAGTCTTCTTTCTACCGTAATCTG
CTGTGGCTGACTGAAAAAGAGGGTTCTTATCCCAAACTGAAGAACTCTTATGTCAACAAG
AAGGGCAAAGAGGTCCTGGTGCTGTGGGGTATCCACCACCCCCCCAACTCCAAGGAGCAG
CAGAATCTGTACCAAAACGAAAATGCTTACGTGTCTGTCGTGACATCTAACTACAACCGG
CGCTTCACGCCCGAAATCGCCGAGCGTCCCAAAGTGCGCGATCAGGCCGGAAGGATGAAC
TACTACTGGACCCTGCTCAAACCCGGAGATACCATCATATTCGAAGCCAATGGCAATCTC
ATCGCTCCCATGTATGCTTTCGCTCTCTCTCGCGGATTCGGTTCCGGCATAATCACTAGC
CAAGCCTCCATGCACGAGTGCAATACCAAGTGTCAAACTCCACTGGGAGCCATTCAAAGC
TCCCTGCCCTACCAAAACATCCATCCTGTGACCATTGGAGAGTGCCCTAAATACGTGCGC
AGCGCGAAACTGCGCATGGTGACCGGTCTCCGCAATATACCCTCCATACAATCGCGCGGC
CTGTTTGGTGCTATCGCTGGCTTTATCGAGGGAGGATGGACTGGAATGATCGACGGCTGG
TATGGCTATCATCATCAAAACGAACAGGGTTCCGGCTACGCCGCTGACCAAAAGTCCACT
CAAAACGCCATTAACGGTATTACAAACAAAGTAAACACCGTGATAGAGAAAATGAATATC
CAATTCACTGCCGTGGGCAAAGAGTTTAACAAGCTGGAGAAGCGCATGGAAAATCTGAAC
AAAAAAGTCGATGATGGCTTCCTCGACATCTGGACTTACAACGCCGAACTCCTCGTGCTG
CTCGAAAACGAGAGGACTCTGGACTTCCACGACTCCAACGTGAAGAACCTGTACGAGAAA
GTCAAATCCCAACTCAAGAACAACGCCAAAGAAATCGGCAACGGCTGCTTCGAGTTCTAC
CACAAATGTGATAACGAGTGTATGGAATCCGTACGGCAAGGCACTTACGACTACCCCAAA
TACAGCGAAGAGAGCAAATTGAACCGCGAGAAAGTGGACGGCGTGAAACTGGAGTCCATG
GGCATCTATCAGCACCACCACCACCACCACTAATAG

SEQ ID NO: 118 ngHA0s

GACACCATCTGCATTGGCTACCATGCCCAGCAGTCTACAGATACAGTCGATACCGTCCTG
GAGAAACAAGTCACCGTCACGCACTCCGTGAACCTCCTGGAGGACTCCCATAATGGCAAA
TTGTGTCGGCTGAAAGGCATCGCCCCTCTGCAACTCGGCAAGTGCAATATTGCGGGGTGG
CTGTTGGGCAACCCAGAATGTGACCCCCTCCTGCCCGTCCGTTCTTGGTCTTACATCGTG
GAAACACCTAACTCCGAGAACGGCATCTGTTACCCGGGCGACTTCATTGACTACGAGGAG
CTCCGCGAACAGCTGTCTTCTGTCTCTTCTTTCGAACGTTTCGAAATTTTCCCAAAGGAA
TCCTCCTGGCCAAACCATAACACCAACGGAGTGACCGCCGCCTGTAGCCATGAGGGCAAG
TCTTCTTTCTACCGTAATCTGCTGTGGCTGACTGAAAAAGAGGGTTCTTATCCCAAACTG
AAGAACTCTTATGTCAACAAGAAGGGCAAAGAGGTCCTGGTGCTGTGGGGTATCCACCAC
CCCCCCAACTCCAAGGAGCAGCAGAATCTGTACCAAAACGAAAATGCTTACGTGTCTGTC
GTGACATCTAACTACAACCGGCGCTTCACGCCCGAAATCGCCGAGCGTCCCAAAGTGCGC
GATCAGGCCGGAAGGATGAACTACTACTGGACCCTGCTCAAACCCGGAGATACCATCATA
TTCGAAGCCAATGGCAATCTCATCGCTCCCATGTATGCTTTCGCTCTCTCTCGCGGATTC
GGTTCCGGCATAATCACTAGCCAAGCCTCCATGCACGAGTGCAATACCAAGTGTCAAACT
CCACTGGGAGCCATTCAAAGCTCCCTGCCCTACCAAAACATCCATCCTGTGACCATTGGA
GAGTGCCCTAAATACGTGCGCAGCGCGAAACTGCGCATGGTGACCGGTCTCCGCAATATA
CCCTCCATACAATCGCGCGGCCTGTTTGGTGCTATCGCTGGCTTTATCGAGGGAGGATGG
ACTGGAATGATCGACGGCTGGTATGGCTATCATCATCAAAACGAACAGGGTTCCGGCTAC
GCCGCTGACCAAAAGTCCACTCAAAACGCCATTAACGGTATTACAAACAAAGTAAACACC
GTGATAGAGAAAATGAATATCCAATTCACTGCCGTGGGCAAAGAGTTTAACAAGCTGGAG
AAGCGCATGGAAAATCTGAACAAAAAAGTCGATGATGGCTTCCTCGACATCTGGACTTAC
AACGCCGAACTCCTCGTGCTGCTCGAAAACGAGAGGACTCTGGACTTCCACGACTCCAAC
GTGAAGAACCTGTACGAGAAAGTCAAATCCCAACTCAAGAACAACGCCAAAGAAATCGGC
AACGGCTGCTTCGAGTTCTACCACAAATGTGATAACGAGTGTATGGAATCCGTACGGCAA

GGCACTTACGACTACCCCAAATACAGCGAAGAGAGCAAATTGAACCGCGAGAAAGTGGAC
GGCGTGAAACTGGAGTCCATGGGCATCTATCAGCACCACCACCACCACCACTAATAG

SEQ ID NO: 119 ngSTF2.HA1-1

ATGAAATTCTTAGTCAACGTTGCCCTTGTTTTTATGGTCGTATACATTTCTTACATCTAT
GCGGCCCAAGTGATCAACACCAACTCCCTGTCCCTGCTCACTCAAAACAACCTCCAGAAA
TCTCAGTCTGCTCTCGGTACTGCTATCGAGCGTCTCTCCTCCGGCCTGAGGATCAACTCC
GCTAAAGATGACGCCGCCGGACAAGCTATCGCTAACCGCTTTACTGCCAACATTAAAGGC
CTGACCCAAGCCTCTAGGAACGCCAATGATGGCATTTCTATCGCTCAAACCACCGAAGGC
GCCCTGAACGAAATCAATAATAACCTGCAACGTGTCCGCGAACTCGCCGTCCAGTCCGCC
CAATCTACCAACTCTCAGTCTGACCTGGACTCTATCCAAGCTGAGATCACCCAAAGGCTC
AACGAGATCGATCGTGTGTCTGGTCAAACTCAGTTTAACGGCGTCAAAGTGCTGGCTCAA
GACAACACACTCACCATCCAAGTAGGAGCCAATGACGGCGAGACAATCGATATCGACCTG
AAGCAGATCAATTCTCAAACTCTGGGCCTGGACTCCCTGAACGTCCAAAAGGCTTACGAC
GTGAAGGACACCGCTGTGACAACCAAAGCCTATGCTAATCAAGGAACAACCCTGGACGTG
TCCGGACTCGATGACGCCGCTATCAAGGCCGCTACTGGCGGCACTCAGGGAACTGCCTCT
GTGACCGGAGGCGCTGTGAAGTTCGACGCCGATAACAACAAATACTTCGTCACTATTGGT
GGTTTCACTGGCGCTGACGCCGCTAAGAACGGTGACTACGAGGTCAACGTCGCCACTGAC
GGAACAGTGACACTGGCCGCTGGCGCTACCAAGACCACAATGCCTGCTGGTGCTACTACT
AAGACAGAGGTGCAGGAACTCAAGGACACCCCTGCCGTGGTGTCCGCCGATGCTAAGAAC
GCTCTCATTGCTGGTGGTGTCGACGCTACGGACGCCAACGGAGCCGAGCTTGTGAAAATG
TCCTACACCGACAAGAACGGAAAGACTATCGAGGGAGGTTACGCTCTGAAGGCTGGCGAT
AAGTACTACGCCGCTGATTATGACGAAGCTACAGGCGCCATTAAAGCTAAAACCACATCT
TATACCGCTGCCGATGGTACTACCAAGACTGCTGCCAATCAGCTGGGTGGAGTCGATGGA
AAAACCGAAGTGGTCACAATCGACGGAAAGACCTATCAAGCCTCCAAGGCTGCTGGCCAC
GACTTCAAGGCCCAACCCGAACTGGCCGAGGCTGCTGCTAAAACCACTGAAAACCCCCTG
CAGAAAATTGACGCTGCCCTGGCCCAAGTGGATGCCCTCCGCTCCGACCTCGGCGCTGTG
CAAAACCGCTTCAACTCCGCTATTACAAACCTCGGCAATACCGTCAATCAGCTCTCTGAA
GCTAGGTCTCGTATTGAGGATTCCGATTACGCTACCGAGGTCTCCCAGATGTCCCGTGCC
CAAATTCTCCAACAAGCCGGCACCTCCGTCCTCGCCCAAGCCAATCAGGTGCCACAGAAT
GTGCTGAGCCTGCTGGCTTCTGGATCTGGATCTGGTTCTTCCCATAATGGCAAATTGTGT
CGGCTGAAAGGCATCGCCCCTCTGCAACTCGGCAAGTGCAATATTGCGGGGTGGCTGTTG
GGCAACCCAGAATGTGACCCCCTCCTGCCCGTCCGTTCTTGGTCTTACATCGTGGAAACA
CCTAACTCCGAGAACGGCATCTGTTACCCGGGCGACTTCATTGACTACGAGGAGCTCCGC
GAACAGCTGTCTTCTGTCTCTTCTTTCGAACGTTTCGAAATTTTCCCAAAGGAATCCTCC
TGGCCAAACCATAACACCAACGGAGTGACCGCCGCCTGTAGCCATGAGGGCAAGTCTTCT
TTCTACCGTAATCTGCTGTGGCTGACTGAAAAAGAGGGTTCTTATCCCAAACTGAAGAAC
TCTTATGTCAACAAGAAGGGCAAAGAGGTCCTGGTGCTGTGGGGTATCCACCACCCCCCC
AACTCCAAGGAGCAGCAGAATCTGTACCAAAACGAAAATGCTTACGTGTCTGTCGTGACA
TCTAACTACAACCGGCGCTTCACGCCCGAAATCGCCGAGCGTCCCAAAGTGCGCGATCAG
GCCGGAAGGATGAACTACTACTGGACCCTGCTCAAACCCGGAGATACCATCATATTCGAA
GCCAATGGCAATCTCATCGCTCCCATGTATGCTTTCGCTCTCTCTCGCGGATTCGGTTCC
GGCATAATCACTAGCCAAGCCTCCATGCACGAGTGCAATACCAAGTGTCAAACTCCACTG
GGAGCCATTCAAAGCTCCCTGCCCTACCAAAACATCCATCCTGTGACCATTGGAGAGTGC
CCTAAATACGTGCGCCACCACCACCACCACCACTAATAG

SEQ ID NO: 120 ng HA1-1

TCCCATAATGGCAAATTGTGTCGGCTGAAAGGCATCGCCCCTCTGCAACTCGGCAAGTGC
AATATTGCGGGGTGGCTGTTGGGCAACCCAGAATGTGACCCCCTCCTGCCCGTCCGTTCT
TGGTCTTACATCGTGGAAACACCTAACTCCGAGAACGGCATCTGTTACCCGGGCGACTTC
ATTGACTACGAGGAGCTCCGCGAACAGCTGTCTTCTGTCTCTTCTTTCGAACGTTTCGAA
ATTTTCCCAAAGGAATCCTCCTGGCCAAACCATAACACCAACGGAGTGACCGCCGCCTGT
AGCCATGAGGGCAAGTCTTCTTTCTACCGTAATCTGCTGTGGCTGACTGAAAAAGAGGGT

TCTTATCCCAAACTGAAGAACTCTTATGTCAACAAGAAGGGCAAAGAGGTCCTGGTGCTG
TGGGGTATCCACCACCCCCCCAACTCCAAGGAGCAGCAGAATCTGTACCAAAACGAAAAT
GCTTACGTGTCTGTCGTGACATCTAACTACAACCGGCGCTTCACGCCCGAAATCGCCGAG
CGTCCCAAAGTGCGCGATCAGGCCGGAAGGATGAACTACTACTGGACCCTGCTCAAACCC
GGAGATACCATCATATTCGAAGCCAATGGCAATCTCATCGCTCCCATGTATGCTTTCGCT
CTCTCTCGCGGATTCGGTTCCGGCATAATCACTAGCCAAGCCTCCATGCACGAGTGCAAT
ACCAAGTGTCAAACTCCACTGGGAGCCATTCAAAGCTCCCTGCCCTACCAAAACATCCAT
CCTGTGACCATTGGAGAGTGCCCTAAATACGTGCGCCACCACCACCACCACCACTAATAG

SEQ ID NO: 121 ngSTF2.HA1-2

ATGAAATTCTTAGTCAACGTTGCCCTTGTTTTTATGGTCGTATACATTTCTTACATCTAT
GCGGCCCAAGTGATCAACACCAACTCCCTGTCCCTGCTCACTCAAAACAACCTCCAGAAA
TCTCAGTCTGCTCTCGGTACTGCTATCGAGCGTCTCTCCTCCGGCCTGAGGATCAACTCC
GCTAAAGATGACGCCGCCGGACAAGCTATCGCTAACCGCTTTACTGCCAACATTAAAGGC
CTGACCCAAGCCTCTAGGAACGCCAATGATGGCATTTCTATCGCTCAAACCACCGAAGGC
GCCCTGAACGAAATCAATAATAACCTGCAACGTGTCCGCGAACTCGCCGTCCAGTCCGCC
CAATCTACCAACTCTCAGTCTGACCTGGACTCTATCCAAGCTGAGATCACCCAAAGGCTC
AACGAGATCGATCGTGTGTCTGGTCAAACTCAGTTTAACGGCGTCAAAGTGCTGGCTCAA
GACAACACACTCACCATCCAAGTAGGAGCCAATGACGGCGAGACAATCGATATCGACCTG
AAGCAGATCAATTCTCAAACTCTGGGCCTGGACTCCCTGAACGTCCAAAAGGCTTACGAC
GTGAAGGACACCGCTGTGACAACCAAAGCCTATGCTAATCAAGGAACAACCCTGGACGTG
TCCGGACTCGATGACGCCGCTATCAAGGCCGCTACTGGCGGCACTCAGGGAACTGCCTCT
GTGACCGGAGGCGCTGTGAAGTTCGACGCCGATAACAACAAATACTTCGTCACTATTGGT
GGTTTCACTGGCGCTGACGCCGCTAAGAACGGTGACTACGAGGTCAACGTCGCCACTGAC
GGAACAGTGACACTGGCCGCTGGCGCTACCAAGACCACAATGCCTGCTGGTGCTACTACT
AAGACAGAGGTGCAGGAACTCAAGGACACCCCTGCCGTGGTGTCCGCCGATGCTAAGAAC
GCTCTCATTGCTGGTGGTGTCGACGCTACGGACGCCAACGGAGCCGAGCTTGTGAAAATG
TCCTACACCGACAAGAACGGAAAGACTATCGAGGGAGGTTACGCTCTGAAGGCTGGCGAT
AAGTACTACGCCGCTGATTATGACGAAGCTACAGGCGCCATTAAAGCTAAAACCACATCT
TATACCGCTGCCGATGGTACTACCAAGACTGCTGCCAATCAGCTGGGTGGAGTCGATGGA
AAAACCGAAGTGGTCACAATCGACGGAAAGACCTATCAAGCCTCCAAGGCTGCTGGCCAC
GACTTCAAGGCCCAACCCGAACTGGCCGAGGCTGCTGCTAAAACCACTGAAAACCCCCTG
CAGAAAATTGACGCTGCCCTGGCCCAAGTGGATGCCCTCCGCTCCGACCTCGGCGCTGTG
CAAAACCGCTTCAACTCCGCTATTACAAACCTCGGCAATACCGTCAATCAGCTCTCTGAA
GCTAGGTCTCGTATTGAGGATTCCGATTACGCTACCGAGGTCTCCCAGATGTCCCGTGCC
CAAATTCTCCAACAAGCCGGCACCTCCGTCCTCGCCCAAGCCAATCAGGTGCCACAGAAT
GTGCTGAGCCTACTAGCATCAGGTTCCGGCTCAGGTTCCAAAGGCATAGCGCCACTGCAG
CTCGGCAAATGTAACATCGCAGGTTGGCTCCTTGGTAACCCGGAGTGCGACCCCCTCCTC
CCTGTACGATCCTGGAGTTATATCGTGGAGACCCCCAATAGCGAGAACGGAATTTGCTAC
CCAGGAGATTTCATAGACTACGAGGAGTTGCGCGAGCAGCTTTCGTCTGTGAGCAGCTTC
GAGAGGTTCGAAATCTTCCCGAAGGAGAGCAGCTGGCCGAATCATAACACTAACGGTGTG
ACAGCCGCCTGCAGTCATGAAGGAAAGAGTTCATTCTATCGCAACCTGCTGTGGTTGACG
GAGAAAGAGGGCAGCTACCCTAAGTTGAAGAACTCCTATGTGAACAAAAAAGGCAAGGAG
GTTCTGGTGCTGTGGGGCATACACCACCCCCCCAATAGCAAGGAGCAGCAGAATCTGTAC
CAAAACGAGAATGCCTATGTGAGCGTGGTCACTAGTAACTATAACCGTCGGTTCACTCCC
GAGATCGCCGAGCGTCCGAAGGTGAGGGACCAGGCAGGCCGGATGAACTACTACTGGACC
CTATTGAAGCCAGGGGACACGATTATCTTCGAGGCAAACGGAAACCTCATAGCGCCGATG
TACGCGTTCGCCCTGAGCCGTGGCTTTGGATCGGGGATCATCACGTCTCACCATCATCAC
CACCATTAATAG

SEQ ID NO: 122 ng HA1-2

AAAGGCATAGCGCCACTGCAGCTCGGCAAATGTAACATCGCAGGTTGGCTCCTTGGTAAC
CCGGAGTGCGACCCCCTCCTCCCTGTACGATCCTGGAGTTATATCGTGGAGACCCCCAAT

AGCGAGAACGGAATTTGCTACCCAGGAGATTTCATAGACTACGAGGAGTTGCGCGAGCAG
CTTTCGTCTGTGAGCAGCTTCGAAAGGTTCGAGATATTCCCGAAGGAGAGCAGCTGGCCG
AATCATAACACTAACGGTGTGACAGCCGCCTGCAGTCATGAAGGAAAGAGTTCATTCTAT
CGCAACCTGCTGTGGTTGACGGAGAAAGAGGGCAGCTACCCTAAGTTGAAGAACTCCTAT
GTGAACAAAAAAGGCAAGGAGGTTCTGGTGCTGTGGGGCATACACCACCCCCCCAATAGC
AAGGAGCAGCAGAATCTGTACCAAAACGAGAATGCCTATGTGAGCGTGGTCACTAGTAAC
TATAACCGTCGGTTCACTCCCGAGATCGCCGAGCGTCCGAAGGTGAGGGACCAGGCAGGC
CGGATGAACTACTACTGGACCCTATTGAAGCCAGGGGACACGATTATCTTCGAGGCAAAC
GGAAACCTCATAGCGCCGATGTACGCCTTCGCCCTGAGCCGTGGCTTTGGATCGGGGATC
ATCACGTCTCACCATCATCACCACCATTAATAG

SEQ ID NO: 123 ngSTF2.HA1-2mut

ATGAAATTCTTAGTCAACGTTGCCCTTGTTTTTATGGTCGTATACATTTCTTACATCTAT
GCGGCCCAAGTGATCAACACCAACTCCCTGTCCCTGCTCACTCAAAACAACCTCCAGAAA
TCTCAGTCTGCTCTCGGTACTGCTATCGAGCGTCTCTCCTCCGGCCTGAGGATCAACTCC
GCTAAAGATGACGCCGCCGGACAAGCTATCGCTAACCGCTTTACTGCCAACATTAAAGGC
CTGACCCAAGCCTCTAGGAACGCCAATGATGGCATTTCTATCGCTCAAACCACCGAAGGC
GCCCTGAACGAAATCAATAATAACCTGCAACGTGTCCGCGAACTCGCCGTCCAGTCCGCC
CAATCTACCAACTCTCAGTCTGACCTGGACTCTATCCAAGCTGAGATCACCCAAAGGCTC
AACGAGATCGATCGTGTGTCTGGTCAAACTCAGTTTAACGGCGTCAAAGTGCTGGCTCAA
GACAACACACTCACCATCCAAGTAGGAGCCAATGACGGCGAGACAATCGATATCGACCTG
AAGCAGATCAATTCTCAAACTCTGGGCCTGGACTCCCTGAACGTCCAAAAGGCTTACGAC
GTGAAGGACACCGCTGTGACAACCAAAGCCTATGCTAATCAAGGAACAACCCTGGACGTG
TCCGGACTCGATGACGCCGCTATCAAGGCCGCTACTGGCGGCACTCAGGGAACTGCCTCT
GTGACCGGAGGCGCTGTGAAGTTCGACGCCGATAACAACAAATACTTCGTCACTATTGGT
GGTTTCACTGGCGCTGACGCCGCTAAGAACGGTGACTACGAGGTCAACGTCGCCACTGAC
GGAACAGTGACACTGGCCGCTGGCGCTACCAAGACCACAATGCCTGCTGGTGCTACTACT
AAGACAGAGGTGCAGGAACTCAAGGACACCCCTGCCGTGGTGTCCGCCGATGCTAAGAAC
GCTCTCATTGCTGGTGGTGTCGACGCTACGGACGCCAACGGAGCCGAGCTTGTGAAAATG
TCCTACACCGACAAGAACGGAAAGACTATCGAGGGAGGTTACGCTCTGAAGGCTGGCGAT
AAGTACTACGCCGCTGATTATGACGAAGCTACAGGCGCCATTAAAGCTAAAACCACATCT
TATACCGCTGCCGATGGTACTACCAAGACTGCTGCCAATCAGCTGGGTGGAGTCGATGGA
AAAACCGAAGTGGTCACAATCGACGGAAAGACCTATCAAGCCTCCAAGGCTGCTGGCCAC
GACTTCAAGGCCCAACCCGAACTGGCCGAGGCTGCTGCTAAAACCACTGAAAACCCCCTG
CAGAAAATTGACGCTGCCCTGGCCCAAGTGGATGCCCTCCGCTCCGACCTCGGCGCTGTG
CAAAACCGCTTCAACTCCGCTATTACAAACCTCGGCAATACCGTCAATCAGCTCTCTGAA
GCTAGGTCTCGTATTGAGGATTCCGATTACGCTACCGAGGTCTCCCAGATGTCCCGTGCC
CAAATTCTCCAACAAGCCGGCACCTCCGTCCTCGCCCAAGCCAATCAGGTGCCACAGAAT
GTGCTGAGCCTGCTGGCTTCTGGATCTGGATCTGGTTCTAAAGGCGCTGCCCCTCTGCAA
CTCGGCAAGTGCAATATTGCGGGGTGGCTGTTGGGCAACCCAGAATGTGACCCCCTCCTG
CCCGTCCGTTCTTGGTCTGACATCGCTGAAACACCTAACTCCGAGAACGGCATCTGTTAC
CCGGGCGACTTCATTGACTACGAGGAGCTCCGCGAACAGCTGTCTTCTGTCTCTTCTTTC
GAACGTTTCGAAATTTTCCCAAAGGAATCCTCCTGGCCAAACCATAACACCAACGGAGTG
ACCGCCGCCTGTAGCCATGAGGGCAAGTCTTCTTTCTACCGTAATCTGCTGTGGCTGACT
GAAAAAGAGGGTTCTTATCCCAAACTGAAGAACTCTTATGTCAACAAGAAGGGCAAAGAG
GTCCTGGTGCTGTGGGGTATCCACCACCCCCCCAACTCCAAGGAGCAGCAGAATCTGTAC
CAAAACGAAAATGCTTACGTGTCTGTCGTGACATCTAACTACAACCGGCGCTTCACGCCC
GAAATCGCCGAGCGTCCCAAAGTGCGCGATCAGGCCGGAAGGATGAACTACTACTGGACC
CTGCTCAAACCCGGAGATACCATCATATTCGAAGCCAATGGCAATCTCATCGCTCCCATG
TATGCTTTCGCTCTCTCTCGCGGATTCGGTTCCGGCATAATCACTAGCCACCACCACCAC
CACCACTAATAG

SEQ ID NO: 124 ngHA1-2*mut*

AAAGGCGCTGCCCCTCTGCAACTCGGCAAGTGCAATATTGCGGGGTGGCTGTTGGGCAAC
CCAGAATGTGACCCCCTCCTGCCCGTCCGTTCTTGGTCTGACATCGCTGAAACACCTAAC
TCCGAGAACGGCATCTGTTACCCGGGCGACTTCATTGACTACGAGGAGCTCCGCGAACAG
CTGTCTTCTGTCTCTTCTTTCGAACGTTTCGAAATTTTCCCAAAGGAATCCTCCTGGCCA
AACCATAACACCAACGGAGTGACCGCCGCCTGTAGCCATGAGGGCAAGTCTTCTTTCTAC
CGTAATCTGCTGTGGCTGACTGAAAAAGAGGGTTCTTATCCCAAACTGAAGAACTCTTAT
GTCAACAAGAAGGGCAAAGAGGTCCTGGTGCTGTGGGGTATCCACCACCCCCCCAACTCC
AAGGAGCAGCAGAATCTGTACCAAAACGAAAATGCTTACGTGTCTGTCGTGACATCTAAC
TACAACCGGCGCTTCACGCCCGAAATCGCCGAGCGTCCCAAAGTGCGCGATCAGGCCGGA
AGGATGAACTACTACTGGACCCTGCTCAAACCCGGAGATACCATCATATTCGAAGCCAAT
GGCAATCTCATCGCTCCCATGTATGCTTTCGCTCTCTCTCGCGGATTCGGTTCCGGCATA
ATCACTAGCCACCACCACCACCACCACTAATAG

SEQ ID NO: 125 ngSTF2.HA1-3

ATGAAATTCTTAGTCAACGTTGCCCTTGTTTTTATGGTCGTATACATTTCTTACATCTAT
GCGGCCCAAGTGATCAACACCAACTCCCTGTCCCTGCTCACTCAAAACAACCTCCAGAAA
TCTCAGTCTGCTCTCGGTACTGCTATCGAGCGTCTCTCCTCCGGCCTGAGGATCAACTCC
GCTAAAGATGACGCCGCCGGACAAGCTATCGCTAACCGCTTTACTGCCAACATTAAAGGC
CTGACCCAAGCCTCTAGGAACGCCAATGATGGCATTTCTATCGCTCAAACCACCGAAGGC
GCCCTGAACGAAATCAATAATAACCTGCAACGTGTCCGCGAACTCGCCGTCCAGTCCGCC
CAATCTACCAACTCTCAGTCTGACCTGGACTCTATCCAAGCTGAGATCACCCAAAGGCTC
AACGAGATCGATCGTGTGTCTGGTCAAACTCAGTTTAACGGCGTCAAAGTGCTGGCTCAA
GACAACACACTCACCATCCAAGTAGGAGCCAATGACGGCGAGACAATCGATATCGACCTG
AAGCAGATCAATTCTCAAACTCTGGGCCTGGACTCCCTGAACGTCCAAAAGGCTTACGAC
GTGAAGGACACCGCTGTGACAACCAAAGCCTATGCTAATCAAGGAACAACCCTGGACGTG
TCCGGACTCGATGACGCCGCTATCAAGGCCGCTACTGGCGGCACTCAGGGAACTGCCTCT
GTGACCGGAGGCGCTGTGAAGTTCGACGCCGATAACAACAAATACTTCGTCACTATTGGT
GGTTTCACTGGCGCTGACGCCGCTAAGAACGGTGACTACGAGGTCAACGTCGCCACTGAC
GGAACAGTGACACTGGCCGCTGGCGCTACCAAGACCACAATGCCTGCTGGTGCTACTACT
AAGACAGAGGTGCAGGAACTCAAGGACACCCCTGCCGTGGTGTCCGCCGATGCTAAGAAC
GCTCTCATTGCTGGTGGTGTCGACGCTACGGACGCCAACGGAGCCGAGCTTGTGAAAATG
TCCTACACCGACAAGAACGGAAAGACTATCGAGGGAGGTTACGCTCTGAAGGCTGGCGAT
AAGTACTACGCCGCTGATTATGACGAAGCTACAGGCGCCATTAAAGCTAAAACCACATCT
TATACCGCTGCCGATGGTACTACCAAGACTGCTGCCAATCAGCTGGGTGGAGTCGATGGA
AAAACCGAAGTGGTCACAATCGACGGAAAGACCTATCAAGCCTCCAAGGCTGCTGGCCAC
GACTTCAAGGCCCAACCCGAACTGGCCGAGGCTGCTGCTAAAACCACTGAAAACCCCCTG
CAGAAAATTGACGCTGCCCTGGCCCAAGTGGATGCCCTCCGCTCCGACCTCGGCGCTGTG
CAAAACCGCTTCAACTCCGCTATTACAAACCTCGGCAATACCGTCAATCAGCTCTCTGAA
GCTAGGTCTCGTATTGAGGATTCCGATTACGCTACCGAGGTCTCCCAGATGTCCCGTGCC
CAAATTCTCCAACAAGCCGGCACCTCCGTCCTCGCCCAAGCCAATCAGGTGCCACAGAAT
GTGCTGAGCCTACTAGCATCAGGTTCCGGCTCAGGTTCCAATAGCGAGAACGGAATTTGC
TACCCAGGAGATTTCATAGACTACGAGGAGTTGCGCGAGCAGCTTTCGTCTGTGAGCAGC
TTCGAGAGGTTCGAAATCTTCCCGAAGGAGAGCAGCTGGCCGAATCATAACACTAACGGT
GTGACAGCCGCCTGCAGTCATGAAGGAAAGAGTTCATTCTATCGCAACCTGCTGTGGTTG
ACGGAGAAGGAGGGCAGCTACCCTAAGCTGAAGAACTCCTATGTGAACAAAAAAGGCAAG
GAGGTTCTGGTGCTGTGGGGCATACACCACCCCCCCAATAGCAAGGAGCAGCAGAATCTG
TACCAAAACGAGAATGCCTATGTGAGCGTGGTCACTAGTAACTATAACCGTCGGTTCACT
CCCGAGATCGCCGAGCGTCCGAAGGTGAGGGACCAGGCAGGCCGGATGAACTACTACTGG
ACCCTATTGAAGCCAGGGGACACGATTATCTTCGAGGCAAACGGAAACCTCATAGCGCCG
ATGTACGCCTTCGCCCTGAGCCGTGGCCACCATCATCACCACCATTAATAG

SEQ ID NO: 126 ngHA1-3

AATAGCGAGAACGGAATTTGCTACCCAGGAGATTTCATAGACTACGAGGAGTTGCGCGAG
CAGCTTTCGTCTGTGAGCAGCTTCGAAAGGTTCGAGATATTCCCGAAGGAGAGCAGCTGG
CCGAATCATAACACTAACGGTGTGACAGCCGCCTGCAGTCATGAAGGAAAGAGTTCATTC
TATCGCAACCTGCTGTGGTTGACGGAGAAAGAGGGCAGCTACCCTAAGTTGAAGAACTCC
TATGTGAACAAAAAAGGCAAGGAGGTTCTGGTGCTGTGGGGCATACACCACCCCCCCAAT
AGCAAGGAGCAGCAGAATCTGTACCAAAACGAGAATGCCTATGTGAGCGTGGTCACTAGT
AACTATAACCGTCGGTTCACTCCCGAGATCGCCGAGCGTCCGAAGGTGAGGGACCAGGCA
GGCCGGATGAACTACTACTGGACCCTATTGAAGCCAGGGGACACGATTATCTTCGAGGCA
AACGGAAACCTCATAGCGCCGATGTACGCCTTCGCCCTGAGCCGTGGCCACCATCATCAC
CACCATTAATAG

SEQ ID NO: 127 ngSTF2.HA1-3*mut*

ATGAAATTCTTAGTCAACGTTGCCCTTGTTTTTATGGTCGTATACATTTCTTACATCTAT
GCGGCCCAAGTGATCAACACCAACTCCCTGTCCCTGCTCACTCAAAACAACCTCCAGAAA
TCTCAGTCTGCTCTCGGTACTGCTATCGAGCGTCTCTCCTCCGGCCTGAGGATCAACTCC
GCTAAAGATGACGCCGCCGGACAAGCTATCGCTAACCGCTTTACTGCCAACATTAAAGGC
CTGACCCAAGCCTCTAGGAACGCCAATGATGGCATTTCTATCGCTCAAACCACCGAAGGC
GCCCTGAACGAAATCAATAATAACCTGCAACGTGTCCGCGAACTCGCCGTCCAGTCCGCC
CAATCTACCAACTCTCAGTCTGACCTGGACTCTATCCAAGCTGAGATCACCCAAAGGCTC
AACGAGATCGATCGTGTGTCTGGTCAAACTCAGTTTAACGGCGTCAAAGTGCTGGCTCAA
GACAACACACTCACCATCCAAGTAGGAGCCAATGACGGCGAGACAATCGATATCGACCTG
AAGCAGATCAATTCTCAAACTCTGGGCCTGGACTCCCTGAACGTCCAAAAGGCTTACGAC
GTGAAGGACACCGCTGTGACAACCAAAGCCTATGCTAATCAAGGAACAACCCTGGACGTG
TCCGGACTCGATGACGCCGCTATCAAGGCCGCTACTGGCGGCACTCAGGGAACTGCCTCT
GTGACCGGAGGCGCTGTGAAGTTCGACGCCGATAACAACAAATACTTCGTCACTATTGGT
GGTTTCACTGGCGCTGACGCCGCTAAGAACGGTGACTACGAGGTCAACGTCGCCACTGAC
GGAACAGTGACACTGGCCGCTGGCGCTACCAAGACCACAATGCCTGCTGGTGCTACTACT
AAGACAGAGGTGCAGGAACTCAAGGACACCCCTGCCGTGGTGTCCGCCGATGCTAAGAAC
GCTCTCATTGCTGGTGGTGTCGACGCTACGGACGCCAACGGAGCCGAGCTTGTGAAAATG
TCCTACACCGACAAGAACGGAAAGACTATCGAGGGAGGTTACGCTCTGAAGGCTGGCGAT
AAGTACTACGCCGCTGATTATGACGAAGCTACAGGCGCCATTAAAGCTAAAACCACATCT
TATACCGCTGCCGATGGTACTACCAAGACTGCTGCCAATCAGCTGGGTGGAGTCGATGGA
AAAACCGAAGTGGTCACAATCGACGGAAAGACCTATCAAGCCTCCAAGGCTGCTGGCCAC
GACTTCAAGGCCCAACCCGAACTGGCCGAGGCTGCTGCTAAAACCACTGAAAACCCCCTG
CAGAAAATTGACGCTGCCCTGGCCCAAGTGGATGCCCTCCGCTCCGACCTCGGCGCTGTG
CAAAACCGCTTCAACTCCGCTATTACAAACCTCGGCAATACCGTCAATCAGCTCTCTGAA
GCTAGGTCTCGTATTGAGGATTCCGATTACGCTACCGAGGTCTCCCAGATGTCCCGTGCC
CAAATTCTCCAACAAGCCGGCACCTCCGTCCTCGCCCAAGCCAATCAGGTGCCACAGAAT
GTGCTGAGCCTGCTGGCTTCTGGATCTGGATCTGGTTCTAACTCCGAGAACGAAATCTGT
TACCCGGGCGACTTCATTGACAAAGAGGAGCTCCGCGAACAGCTGTCTTCTGTCTCTTCT
TTCGAACGTTTCGAAATTTTCCCAAAGGAATCCTCCTGGCCAAACCATAACACCAACGGA
GTGACCGCCGCCTGTAGCCATGAGGGCAAGTCTTCTTTCTACCGTAATCTGCTGTGGCTG
ACTGAAAAAGAGGGTTCTTATCCCAAACTGAAGAACTCTTATGTCAACAAGAAGGGCAAA
GAGGTCCTGGTGCTGTGGGGTATCCACCACCCCCCCAACTCCAAGGAGCAGCAGAATCTG
TACCAAAACGAAAATGCTTACGTGTCTGTCGTGACATCTAACTACAACCGGCGCTTCACG
CCCGAAATCGCCGAGCGTCCCAAAGTGCGCGATCAGGCCGGAAGGATGAACTACTACTGG
ACCCTGCTCAAACCCGGAGATACCATCATATTCGAAGCCAATGGCAATCTCATCGCTCCC
ATGTATGCTGCTGCTCTCTCTCGCGGACACCACCACCACCACCACTAATAG

SEQ ID NO: 128 ngHA1-3*mut*

AACTCCGAGAACGAAATCTGTTACCCGGGCGACTTCATTGACAAAGAGGAGCTCCGCGAA
CAGCTGTCTTCTGTCTCTTCTTTCGAACGTTTCGAAATTTTCCCAAAGGAATCCTCCTGG

CCAAACCATAACACCAACGGAGTGACCGCCGCCTGTAGCCATGAGGGCAAGTCTTCTTTC
TACCGTAATCTGCTGTGGCTGACTGAAAAAGAGGGTTCTTATCCCAAACTGAAGAACTCT
TATGTCAACAAGAAGGGCAAAGAGGTCCTGGTGCTGTGGGGTATCCACCACCCCCCCAAC
TCCAAGGAGCAGCAGAATCTGTACCAAAACGAAAATGCTTACGTGTCTGTCGTGACATCT
AACTACAACCGGCGCTTCACGCCCGAAATCGCCGAGCGTCCCAAAGTGCGCGATCAGGCC
GGAAGGATGAACTACTACTGGACCCTGCTCAAACCCGGAGATACCATCATATTCGAAGCC
AATGGCAATCTCATCGCTCCCATGTATGCTGCTGCTCTCTCTCGCGGACACCACCACCAC
CACCACTAATAG

SEQ ID NO: 129 wtSTF2.HA1-1ng

ATGAAATTCTTAGTCAACGTTGCCCTTGTTTTTATGGTCGTATACATTTCTTACATCTAT
GCGGCCCAGGTTATCAATACCAACTCCCTGTCGTTGCTCACCCAAAATAACCTTAATAAA
AGCCAGAGCGCACTGGGAACCGCCATAGAACGCCTCTCAAGCGGCCTCCGGATCAATTCT
GCAAAAGACGACGCCGCCGGTCAGGCCATCGCAAACCGCTTTACCGCCAATATCAAGGGA
CTGACGCAGGCTTCGAGGAATGCTAACGATGGAATAAGCATCGCTCAAACCACGGAGGGC
GCCCTGAACGAGATCAACAACAACCTACAGCGCGTCAGGGAGCTCGCAGTGCAGTCCGCC
AATTCGACCAACTCGCAGTCGGACCTGGACTCGATCCAAGCCGAAATCACCCAGCGCCTG
AATGAGATTGACCGGGTGAGCGGTCAGACACAGTTTAACGGCGTGAAGGTACTTGCACAG
GATAACACACTTACGATACAGGTGGGCGCCAACGATGGTGAAACCATAGACATTGATCTC
AAACAGATTAACAGCCAGACGCTCGGGTTGGATAGCCTGAATGTGCAAAAGGCGTACGAC
GTGAAAGACACGGCGGTCACTACCAAAGCCTACGCTAACAATGGCACTACCTTGGATGTG
AGCGGATTGGATGATGCAGCAATCAAGGCTGCTACCGGCGGTACGAACGGAACCGCGTCC
GTGACCGGCGGTGCCGTGAAGTTCGATGCTGACAACAATAAGTATTTCGTCACCATTGGA
GGCTTTACTGGCGCCGACGCAGCAAAGAACGGCGACTATGAAGTGAACGTGGCAACCGAT
GGAACCGTGACGCTGGCCGCTGGTGCCACCAAGACCACCATGCCAGCCGGCGCCACAACT
AAGACCGAGGTGCAGGAGTTAAAGGACACCCCCGCGGTGGTTAGCGCAGATGCCAAAAAC
GCCTTGATCGCCGGCGGAGTGGATGCAACTGATGCTAATGGTGCGGAGCTGGTTAAAATG
TCGTATACAGACAAGAATGGTAAGACGATCGAGGGCGGTTATGCCCTTAAGGCAGGAGAT
AAGTATTACGCTGCTGATTACGATGAGGCGACGGGAGCTATTAAGGCCAAGACAACGTCA
TACACGGCGGCGGACGGAACGACTAAGACGGCTGCCAATCAGTTGGGAGGGGTTGACGGG
AAGACAGAGGTCGTTACGATCGATGGCAAGACATACAACGCTTCCAAGGCCGCTGGCCAC
GATTTCAAAGCTCAACCCGAACTGGCCGAGGCCGCGGCGAAAACAACTGAGAACCCGTTG
CAGAAGATTGATGCGGCCCTGGCGCAAGTAGATGCCCTGCGCTCAGACCTGGGCGCCGTT
CAAAATCGATTCAATTCCGCGATTACAAACCTGGGCAATACAGTAAACAATCTATCCGAG
GCCAGATCCCGCATTGAAGACTCCGACTACGCGACAGAAGTAAGTAACATGAGTCGTGCC
CAGATTCTGCAGCAGGCCGGCACTAGTGTCCTGGCCCAGGCCAATCAAGTCCCGCAGAAT
GTGCTGAGCCTACTAGCATCTGGATCTGGATCTGGTTCTTCCCATAATGGCAAATTGTGT
CGGCTGAAAGGCATCGCCCCTCTGCAACTCGGCAAGTGCAATATTGCGGGGTGGCTGTTG
GGCAACCCAGAATGTGACCCCCTCCTGCCCGTCCGTTCTTGGTCTTACATCGTGGAAACA
CCTAACTCCGAGAACGGCATCTGTTACCCGGGCGACTTCATTGACTACGAGGAGCTCCGC
GAACAGCTGTCTTCTGTCTCTTCTTTCGAACGTTTCGAAATTTTCCCAAAGGAATCCTCC
TGGCCAAACCATAACACCAACGGAGTGACCGCCGCCTGTAGCCATGAGGGCAAGTCTTCT
TTCTACCGTAATCTGCTGTGGCTGACTGAAAAAGAGGGTTCTTATCCCAAACTGAAGAAC
TCTTATGTCAACAAGAAGGGCAAAGAGGTCCTGGTGCTGTGGGGTATCCACCACCCCCCC
AACTCCAAGGAGCAGCAGAATCTGTACCAAAACGAAAATGCTTACGTGTCTGTCGTGACA
TCTAACTACAACCGGCGCTTCACGCCCGAAATCGCCGAGCGTCCCAAAGTGCGCGATCAG
GCCGGAAGGATGAACTACTACTGGACCCTGCTCAAACCCGGAGATACCATCATATTCGAA
GCCAATGGCAATCTCATCGCTCCCATGTATGCTTTCGCTCTCTCTCGCGGATTCGGTTCC
GGCATAATCACTAGCCAAGCCTCCATGCACGAGTGCAATACCAAGTGTCAAACTCCACTG
GGAGCCATTCAAAGCTCCCTGCCCTACCAAAACATCCATCCTGTGACCATTGGAGAGTGC
CCTAAATACGTGCGCCACCACCACCACCACCACTAATAG

SEQ ID NO: 130 ngHA1-1

TCCCATAATGGCAAATTGTGTCGGCTGAAAGGCATCGCCCCTCTGCAACTCGGCAAGTGC
AATATTGCGGGGTGGCTGTTGGGCAACCCAGAATGTGACCCCCTCCTGCCCGTCCGTTCT
TGGTCTTACATCGTGGAAACACCTAACTCCGAGAACGGCATCTGTTACCCGGGCGACTTC
ATTGACTACGAGGAGCTCCGCGAACAGCTGTCTTCTGTCTCTTCTTTCGAACGTTTCGAA
ATTTTCCCAAAGGAATCCTCCTGGCCAAACCATAACACCAACGGAGTGACCGCCGCCTGT
AGCCATGAGGGCAAGTCTTCTTTCTACCGTAATCTGCTGTGGCTGACTGAAAAAGAGGGT
TCTTATCCCAAACTGAAGAACTCTTATGTCAACAAGAAGGGCAAAGAGGTCCTGGTGCTG
TGGGGTATCCACCACCCCCCCAACTCCAAGGAGCAGCAGAATCTGTACCAAAACGAAAAT
GCTTACGTGTCTGTCGTGACATCTAACTACAACCGGCGCTTCACGCCCGAAATCGCCGAG
CGTCCCAAAGTGCGCGATCAGGCCGGAAGGATGAACTACTACTGGACCCTGCTCAAACCC
GGAGATACCATCATATTCGAAGCCAATGGCAATCTCATCGCTCCCATGTATGCTTTCGCT
CTCTCTCGCGGATTCGGTTCCGGCATAATCACTAGCCAAGCCTCCATGCACGAGTGCAAT
ACCAAGTGTCAAACTCCACTGGGAGCCATTCAAAGCTCCCTGCCCTACCAAAACATCCAT
CCTGTGACCATTGGAGAGTGCCCTAAATACGTGCGCCACCACCACCACCACCACTAATAG

SEQ ID NO: 131 ngSTF2.HA1-1wt

ATGAAATTCTTAGTCAACGTTGCCCTTGTTTTTATGGTCGTATACATTTCTTACATCTAT
GCGGCCCAAGTGATCAACACCAACTCCCTGTCCCTGCTCACTCAAAACAACCTCCAGAAA
TCTCAGTCTGCTCTCGGTACTGCTATCGAGCGTCTCTCCTCCGGCCTGAGGATCAACTCC
GCTAAAGATGACGCCGCCGGACAAGCTATCGCTAACCGCTTTACTGCCAACATTAAAGGC
CTGACCCAAGCCTCTAGGAACGCCAATGATGGCATTTCTATCGCTCAAACCACCGAAGGC
GCCCTGAACGAAATCAATAATAACCTGCAACGTGTCCGCGAACTCGCCGTCCAGTCCGCC
CAATCTACCAACTCTCAGTCTGACCTGGACTCTATCCAAGCTGAGATCACCCAAAGGCTC
AACGAGATCGATCGTGTGTCTGGTCAAACTCAGTTTAACGGCGTCAAAGTGCTGGCTCAA
GACAACACACTCACCATCCAAGTAGGAGCCAATGACGGCGAGACAATCGATATCGACCTG
AAGCAGATCAATTCTCAAACTCTGGGCCTGGACTCCCTGAACGTCCAAAAGGCTTACGAC
GTGAAGGACACCGCTGTGACAACCAAAGCCTATGCTAATCAAGGAACAACCCTGGACGTG
TCCGGACTCGATGACGCCGCTATCAAGGCCGCTACTGGCGGCACTCAGGGAACTGCCTCT
GTGACCGGAGGCGCTGTGAAGTTCGACGCCGATAACAACAAATACTTCGTCACTATTGGT
GGTTTCACTGGCGCTGACGCCGCTAAGAACGGTGACTACGAGGTCAACGTCGCCACTGAC
GGAACAGTGACACTGGCCGCTGGCGCTACCAAGACCACAATGCCTGCTGGTGCTACTACT
AAGACAGAGGTGCAGGAACTCAAGGACACCCCTGCCGTGGTGTCCGCCGATGCTAAGAAC
GCTCTCATTGCTGGTGGTGTCGACGCTACGGACGCCAACGGAGCCGAGCTTGTGAAAATG
TCCTACACCGACAAGAACGGAAAGACTATCGAGGGAGGTTACGCTCTGAAGGCTGGCGAT
AAGTACTACGCCGCTGATTATGACGAAGCTACAGGCGCCATTAAAGCTAAAACCACATCT
TATACCGCTGCCGATGGTACTACCAAGACTGCTGCCAATCAGCTGGGTGGAGTCGATGGA
AAAACCGAAGTGGTCACAATCGACGGAAAGACCTATCAAGCCTCCAAGGCTGCTGGCCAC
GACTTCAAGGCCCAACCCGAACTGGCCGAGGCTGCTGCTAAAACCACTGAAAACCCCCTG
CAGAAAATTGACGCTGCCCTGGCCCAAGTGGATGCCCTCCGCTCCGACCTCGGCGCTGTG
CAAAACCGCTTCAACTCCGCTATTACAAACCTCGGCAATACCGTCAATCAGCTCTCTGAA
GCTAGGTCTCGTATTGAGGATTCCGATTACGCTACCGAGGTCTCCCAGATGTCCCGTGCC
CAAATTCTCCAACAAGCCGGCACCTCCGTCCTCGCCCAAGCCAATCAGGTGCCACAGAAT
GTGCTGAGCCTACTAGCATCAGGTTCCGGCTCAGGTTCCAGCCACAACGGAAAGCTGTGC
CGTCTGAAAGGCATAGCGCCACTGCAGCTCGGCAAATGTAACATCGCAGGTTGGCTCCTT
GGTAACCCGGAGTGCGACCCCCTCCTCCCTGTACGATCCTGGAGTTATATCGTGGAGACC
CCCAATAGCGAGAACGGAATTTGCTACCCAGGAGATTTCATAGACTACGAGGAGTTGCGC
GAGCAGCTTTCGTCTGTGAGCAGCTTCGAGAGGTTCGAAATCTTCCCGAAGGAGAGCAGC
TGGCCGAATCATAACACTAACGGTGTGACAGCCGCCTGCAGTCATGAAGGAAAGAGTTCA
TTCTATCGCAACCTGCTGTGGTTGACGGAGAAAGAGGGCAGCTACCCTAAGTTGAAGAAC
TCCTATGTGAACAAAAAAGGCAAGGAGGTTCTGGTGCTGTGGGGCATACACCACCCCCCC
AATAGCAAGGAGCAGCAGAATCTGTACCAAAACGAGAATGCCTATGTGAGCGTGGTCACT
AGTAACTATAACCGTCGGTTCACTCCCGAGATCGCCGAGCGTCCGAAGGTGAGGGACCAG

GCAGGCCGGATGAACTACTACTGGACCCTATTGAAGCCAGGGGACACGATTATCTTCGAG
GCAAACGGAAACCTCATAGCGCCGATGTACGCGTTCGCCCTGAGCCGCGGCTTTGGATCG
GGGATCATCACGTCTAACGCCTCGATGCACGAATGTAATACCAAATGCCAGACCCCACTG
GGTGCTATCAACTCGTCCTTACCCTATCAAAATATACATCCGGTCACCATAGGCGAGTGT
CCCAAATATGTCAGACACCATCATCACCACCATTAATAG

SEQ ID NO: 132 HA1-1 PR8
AGCCACAACGGAAAGCTGTGCCGTCTGAAAGGCATAGCGCCACTGCAGCTCGGCAAATGT
AACATCGCAGGTTGGCTCCTTGGTAACCCGGAGTGCGACCCCCTCCTCCCTGTACGATCC
TGGAGTTATATCGTGGAGACCCCCAATAGCGAGAACGGAATTTGCTACCCAGGAGATTTC
ATAGACTACGAGGAGTTGCGCGAGCAGCTTTCGTCTGTGAGCAGCTTCGAAAGGTTCGAG
ATATTCCCGAAGGAGAGCAGCTGGCCGAATCATAACACTAACGGTGTGACAGCCGCCTGC
AGTCATGAAGGAAAGAGTTCATTCTATCGCAACCTGCTGTGGTTGACGGAGAAAGAGGGC
AGCTACCCTAAGTTGAAGAACTCCTATGTGAACAAAAAAGGCAAGGAGGTTCTGGTGCTG
TGGGGCATACACCACCCCCCCAATAGCAAGGAGCAGCAGAATCTGTACCAAAACGAGAAT
GCCTATGTGAGCGTGGTCACTAGTAACTATAACCGTCGGTTCACTCCCGAGATCGCCGAG
CGTCCGAAGGTGAGGGACCAGGCAGGCCGGATGAACTACTACTGGACCCTATTGAAGCCA
GGGGACACGATTATCTTCGAGGCAAACGGAAACCTCATAGCGCCGATGTACGCCTTCGCC
CTGAGCCGTGGCTTTGGATCGGGGATCATCACGTCTAACGCCTCGATGCACGAATGTAAT
ACCAAATGCCAGACCCCACTGGGTGCTATCAACTCGTCCTTACCCTATCAAAATATACAT
CCGGTCACCATAGGCGAGTGTCCCAAATATGTCAGACACCATCATCACCACCATTAATAG

SEQ ID NO: 133 HA1-1 PR8
ATGAAATTCTTAGTCAACGTTGCCCTTGTTTTTATGGTCGTATACATTTCTTACATCTAT
GCGAGCCACAACGGAAAGCTGTGCCGTCTGAAAGGCATAGCGCCACTGCAGCTCGGCAAA
TGTAACATCGCAGGTTGGCTCCTTGGTAACCCGGAGTGCGACCCCCTCCTCCCTGTACGA
TCCTGGAGTTATATCGTGGAGACCCCCAATAGCGAGAACGGAATTTGCTACCCAGGAGAT
TTCATAGACTACGAGGAGTTGCGCGAGCAGCTTTCGTCTGTGAGCAGCTTCGAAAGGTTC
GAGATATTCCCGAAGGAGAGCAGCTGGCCGAATCATAACACTAACGGTGTGACAGCCGCC
TGCAGTCATGAAGGAAAGAGTTCATTCTATCGCAACCTGCTGTGGTTGACGGAGAAAGAG
GGCAGCTACCCTAAGTTGAAGAACTCCTATGTGAACAAAAAAGGCAAGGAGGTTCTGGTG
CTGTGGGGCATACACCACCCCCCCAATAGCAAGGAGCAGCAGAATCTGTACCAAAACGAG
AATGCCTATGTGAGCGTGGTCACTAGTAACTATAACCGTCGGTTCACTCCCGAGATCGCC
GAGCGTCCGAAGGTGAGGGACCAGGCAGGCCGGATGAACTACTACTGGACCCTATTGAAG
CCAGGGGACACGATTATCTTCGAGGCAAACGGAAACCTCATAGCGCCGATGTACGCCTTC
GCCCTGAGCCGTGGCTTTGGATCGGGGATCATCACGTCTAACGCCTCGATGCACGAATGT
AATACCAAATGCCAGACCCCACTGGGTGCTATCAACTCGTCCTTACCCTATCAAAATATA
CATCCGGTCACCATAGGCGAGTGTCCCAAATATGTCAGACACCATCATCACCACCATTAAT
AG

SEQ ID NO: 134 FOR PRIMER
AGGCAGATCTATGAAATTCTTAGTCAACGTTGCCCTTGTTTTTATGGTCGTATACATTTC
TTACATCTATGCGAGCCACAACGGAAAGCTGTGCCGTCTGAAAGG

SEQ ID NO: 135 REV PRIMER
ACCTGCATGCCTATTAATGGTGGTGATGATGGTGTCTGACATATTTGGGACACTC

SEQ ID NO: 136 HA0s PR8
ATGAAATTCTTAGTCAACGTTGCCCTTGTTTTTATGGTCGTATACATTTCTTACATCTAT
GCGGACACCATTTGCATTGGATACCATGCAAACAACTCAACCGATACTGTTGATACCGTC
CTTGAGAAGAACGTTACCGTCACGCACTCGGTCAACCTATTAGAGGATAGCCACAACGGA
AAGCTGTGCCGTCTGAAAGGCATAGCGCCACTGCAGCTCGGCAAATGTAACATCGCAGGT
TGGCTCCTTGGTAACCCGGAGTGCGACCCCCTCCTCCCTGTACGATCCTGGAGTTATATC

GTGGAGACCCCCAATAGCGAGAACGGAATTTGCTACCCAGGAGATTTCATAGACTACGAG
GAGTTGCGCGAGCAGCTTTCGTCTGTGAGCAGCTTCGAAAGGTTCGAGATATTCCCGAAG
GAGAGCAGCTGGCCGAATCATAACACTAACGGTGTGACAGCCGCCTGCAGTCATGAAGGA
AAGAGTTCATTCTATCGCAACCTGCTGTGGTTGACGGAGAAAGAGGGCAGCTACCCTAAG
TTGAAGAACTCCTATGTGAACAAAAAAGGCAAGGAGGTTCTGGTGCTGTGGGGCATACAC
CACCCCCCCAATAGCAAGGAGCAGCAGAATCTGTACCAAAACGAGAATGCCTATGTGAGC
GTGGTCACTAGTAACTATAACCGTCGGTTCACTCCCGAGATCGCCGAGCGTCCGAAGGTG
AGGGACCAGGCAGGCCGGATGAACTACTACTGGACCCTATTGAAGCCAGGGGACACGATT
ATCTTCGAGGCAAACGGAAACCTCATAGCGCCGATGTACGCCTTCGCCCTGAGCCGTGGC
TTTGGATCGGGGATCATCACGTCTAACGCCTCGATGCACGAATGTAATACCAAATGCCAG
ACCCCACTGGGTGCTATCAACTCGTCCTTACCCTATCAAAATATACATCCGGTCACCATA
GGCGAGTGTCCCAAATATGTCAGATCCGCCAAGTTGCGGATGGTGACCGGCCTCCGCAAT
ATTCCTAGTATTCAGTCACGCGGCTTGTTCGGCGCCATCGCTGGTTTCATCGAAGGCGGG
TGGACAGGCATGATTGATGGCTGGTATGGCTATCACCACCAGAACGAGCAGGGCTCGGGC
TACGCGGCTGACCAGAAGTCGACTCAGAATGCCATCAATGGCATCACGAATAAGGTGAAC
ACGGTCATTGAAAAGATGAACATTCAATTTACAGCCGTAGGAAAAGAGTTTAATAAACTG
GAAAAAAGAATGGAGAATCTGAATAAGAAGGTGGACGACGGATTTTTGGACATCTGGACG
TACAACGCCGAGCTGCTGGTTCTGCTGGAAAATGAGCGAACACTGGATTTTCATGATTCT
AACGTAAAGAATCTGTACGAGAAGGTGAAGTCCCAACTAAAGAATAATGCCAAGGAAATC
GGAAATGGATGCTTTGAGTTTTACCACAAGTGCGATAATGAGTGCATGGAATCCGTGCGA
AATGGTACATACGATTACCCAAAGTACTCCGAAGAATCCAAGCTAAATCGCGAAAAGGTT
GATGGTGTTAAACTTGAATCCATGGGTATTTACCAACACCATCATCACCACCATTAATAG

SEQ ID NO: 137 HA1-1 PR8 (no his tag)

ATGAAATTCTTAGTCAACGTTGCCCTTGTTTTTATGGTCGTATACATTTCTTACATCTAT
GCGAGCCACAACGGAAAGCTGTGCCGTCTGAAAGGCATAGCGCCACTGCAGCTCGGCAAA
TGTAACATCGCAGGTTGGCTCCTTGGTAACCCGGAGTGCGACCCCCTCCTCCCTGTACGA
TCCTGGAGTTATATCGTGGAGACCCCCAATAGCGAGAACGGAATTTGCTACCCAGGAGAT
TTCATAGACTACGAGGAGTTGCGCGAGCAGCTTTCGTCTGTGAGCAGCTTCGAAAGGTTC
GAGATATTCCCGAAGGAGAGCAGCTGGCCGAATCATAACACTAACGGTGTGACAGCCGCC
TGCAGTCATGAAGGAAAGAGTTCATTCTATCGCAACCTGCTGTGGTTGACGGAGAAAGAG
GGCAGCTACCCTAAGTTGAAGAACTCCTATGTGAACAAAAAAGGCAAGGAGGTTCTGGTG
CTGTGGGGCATACACCACCCCCCCAATAGCAAGGAGCAGCAGAATCTGTACCAAAACGAG
AATGCCTATGTGAGCGTGGTCACTAGTAACTATAACCGTCGGTTCACTCCCGAGATCGCC
GAGCGTCCGAAGGTGAGGGACCAGGCAGGCCGGATGAACTACTACTGGACCCTATTGAAG
CCAGGGGACACGATTATCTTCGAGGCAAACGGAAACCTCATAGCGCCGATGTACGCCTTC
GCCCTGAGCCGTGGCTTTGGATCGGGGATCATCACGTCTAACGCCTCGATGCACGAATGT
AATACCAAATGCCAGACCCCACTGGGTGCTATCAACTCGTCCTTACCCTATCAAAATATA
CATCCGGTCACCATAGGCGAGTGTCCCAAATATGTCAGATAATAG

SEQ ID NO: 138 REV PRIMER

ACCTGCATGCCTATTATCTGACATATTTGGGACACTCGCCTATGG

SEQ ID NO: 139 HA1-1 NC

ATGAAATTCTTAGTCAACGTTGCCCTTGTTTTTATGGTCGTATACATTTCTTACATCTAT
GCGAGCCACAATGGTAAGCTGTGTCTGTTGAAGGGAATTGCACCCCTCCAACTGGGCAAC
TGTTCGGTCGCTGGTTGGATCCTCGGAAACCCAGAATGCGAGCTCCTTATCAGTAAGGAA
TCTTGGTCTTATATTGTCGAAACCCCGAACCCCGAGAACGGAACATGCTACCCGGGTTAC
TTTGCTGATTACGAAGAGCTTCGCGAGCAACTCAGCTCCGTATCCTCCTTCGAGCGCTTC
GAGATTTTTCCCAAAGAGTCCAGCTGGCCAAATCATACCGTCACCGGCGTGTCGGCCTCC
TGTTCCCACAACGGAAAGTCTAGCTTCTATAGAAATCTTCTCTGGCTGACGGGTAAGAAT
GGTCTTTACCCCAATTTGAGCAAGTCCTACGTCAACAACAAAGAAAAGGAAGTTCTGGTA
TTGTGGGGTGTGCACCACCCTCCGAACATCGGCAATCAGCGCGCCCTGTATCACACAGAG

AACGCGTATGTTTCCGTTGTCTCCTCACATTACTCGAGGCGCTTCACTCCTGAAATAGCT
AAGCGTCCGAAAGTGCGTGACCAGGAGGGACGTATCAACTATTATTGGACGCTGTTGGAG
CCAGGCGATACAATTATCTTCGAGGCTAACGGTAACCTTATCGCTCCCTGGTACGCCTTC
GCCCTGTCGCGTGGTTTCGGTAGTGGAATAATCACTAGTAATGCTCCTATGGACGAGTGT
GACGCTAAGTGCCAAACACCTCAGGGCGCTATCAATAGCTCCCTTCCATTCCAGAACGTC
CATCCGGTTACCATTGGAGAGTGTCCAAAGTACGTGAGACACCATCATCACCATCACTAA
TAG

SEQ ID NO: 140 FOR PRIMER

AGGCAGATCTATGAAATTCTTAGTCAACGTTGCCCTTGTTTTTATGGTCGTATACATTTC
TTACATCTATGCGAGCCACAATGGTAAGCTGTGTCTGTTGAAG

SEQ ID NO: 141 REV PRIMER

AGTCGCATGCCTATTAATGGTGGTGATGATGGTGTCTCACGTACTTTGGACACTCTCCAAT
GG

SEQ ID NO: 142 HA0s NC

AGATCTATGAAGTTTCTCGTGAACGTTGCACTGGTGTTTATGGTAGTCTATATATCGTAC
ATTTATGCTGACACGATTTGCATCGGTTACCATGCGAACAACTCCACGGATACCGTGGAC
ACAGTGTTGGAAAAGAACGTGACCGTCACGCACTCCGTAAACCTTCTGGAGGACAGCCAC
AATGGTAAGCTGTGTCTGTTGAAGGGAATTGCACCCCTCCAACTGGGCAACTGTTCGGTC
GCTGGTTGGATCCTCGGAAACCCAGAATGCGAGCTCCTTATCAGTAAGGAATCTTGGTCT
TATATTGTCGAAACCCCGAACCCCGAGAACGGAACATGCTACCCGGGTTACTTTGCTGAT
TACGAAGAGCTTCGCGAGCAACTCAGCTCCGTATCCTCCTTCGAGCGCTTCGAGATTTTT
CCCAAAGAGTCCAGCTGGCCAAATCATACCGTCACCGGCGTGTCGGCCTCCTGTTCCCAC
AACGGAAAGTCTAGCTTCTATAGAAATCTTCTCTGGCTGACGGGTAAGAATGGTCTTTAC
CCCAATTTGAGCAAGTCCTACGTCAACAACAAAGAAAAGGAAGTTCTGGTATTGTGGGGT
GTGCACCACCCTCCGAACATCGGCAATCAGCGCGCCCTGTATCACACAGAGAACGCGTAT
GTTTCCGTTGTCTCCTCACATTACTCGAGGCGCTTCACTCCTGAAATAGCTAAGCGTCCG
AAAGTGCGTGACCAGGAGGGACGTATCAACTATTATTGGACGCTGTTGGAGCCAGGCGAT
ACAATTATCTTCGAGGCTAACGGTAACCTTATCGCTCCCTGGTACGCCTTCGCCCTGTCG
CGTGGTTTCGGTAGTGGAATAATCACTAGTAATGCTCCTATGGACGAGTGTGACGCTAAG
TGCCAAACACCTCAGGGCGCTATCAATAGCTCCCTTCCATTCCAGAACGTCCATCCGGTT
ACCATTGGAGAGTGTCCAAAGTACGTGAGATCGGCCAAACTTCGCATGGTCACGGGTCTG
CGCAACATCCCGTCAATCCAATCTAGGGGCCTCTTCGGCGCTATCGCCGGTTTCATTGAG
GGCGGTTGGACTGGAATGGTTGACGGATGGTACGGCTATCATCACCAGAACGAACAAGGT
TCCGGTTACGCTGCTGACCAGAAATCTACTCAGAACGCGATCAATGGTATCACGAACAAG
GTGAACAGCGTCATTGAAAAGATGAATACTCAGTTTACAGCCGTGGGCAAAGAGTTCAAT
AAACTCGAGAGACGTATGGAAAACCTCAATAAGAAGGTGGATGACGGCTTCCTGGACATT
TGGACTTACAACGCCGAGCTGCTGGTCCTGCTCGAGAACGAGAGAACCCTTGACTTCCAC
GACAGCAACGTCAAGAACCTGTACGAGAAGGTGAAAAGTCAACTTAAAAACAATGCCAAG
GAGATTGGTAACGGCTGCTTCGAATTCTACCACAAGTGTAATAATGAGTGCATGGAATCC
GTTAAGAACGGCACCTACGATTACCCTAAATACTCAGAGGAGTCCAAGCTTAACCGCGAG
AAGATCGACGGCGTAAAACTGGAAAGCATGGGCGTATACCAACACCATCATCACCATCAC
TAATAGGCATGC

SEQ ID NO: 143 HA1-1 VN

ATGAAATTCTTAGTCAACGTTGCCCTTGTTTTTATGGTCGTATACATTTCTTACATCTAT
GCGGAGAAGAAACACAACGGTAAGCTTTGCGACTTGGATGGAGTCAAGCCCCTCATACTT
AGAGATTGTAGTGTAGCCGGTTGGCTGCTCGGTAACCCAATGTGCGATGAGTTCATCAAT
GTTCCCGAATGGTCATATATCGTCGAAAAAGCTAATCCTGTCAACGACCTGTGCTACCCC
GGTGATTTCAATGACTATGAAGAACTGAAGCACCTGCTCTCCCGCATCAACCATTTCGAG
AAAATCCAGATCATTCCCAAGAGTTCCTGGTCTAGCCATGAGGCTAGTCTGGGTGTCTCA

TCCGCCTGCCCATATCAGGGTAAAAGTTCTTTCTTTAGGAACGTAGTATGGTTGATAAAG
AAAAACTCTACATACCCGACCATCAAGCGCTCTTACAACAATACGAACCAAGAGGATCTG
CTTGTCCTTTGGGGAATCCATCATCCTAATGATGCTGCCGAACAGACTAAGCTCTACCAA
AACCCTACCACTTATATTTCCGTCGGCACCTCTACTCTGAACCAGCGCCTTGTGCCCAGG
ATCGCTACGAGATCAAAAGTCAACGGCCAATCGGGCCGCATGGAATTCTTCTGGACGATC
CTGAAGCCTAATGACGCTATCAACTTCGAGTCAAATGGAAACTTTATCGCTCCCGAGTAC
GCTTACAAGATCGTCAAGAAGGGCGACTCCACGATTATGAAGTCAGAGTTGGAGTACGGC
AACTGCAACACAAAGTGCCAAACTCCTATGGGCGCTATAAATTCTTCAATGCCGTTCCAC
AACATCCATCCGCTCACGATCGGTGAGTGCCCGAAATATGTAAAGCACCATCACCACCAT
CACTAATAG

SEQ ID NO: 144 FOR PRIMER
AGGCAGATCTATGAAATTCTTAGTCAACGTTGCCCTTGTTTTTATGGTCGTATACATTTC
TTACATCTATGCGGAGAAGAAACACAACGGTAAGCTTTG

SEQ ID NO: 145 REV PRIMER
AGTCGCATGCCTATTATGGTGGTGATGATGGTGCTTTACATATTTCGGGCACTCACCG

SEQ ID NO: 146 HA0s VN
AGATCTATGAAATTCTTGGTTAATGTAGCCCTGGTGTTTATGGTAGTGTACATTTCATAC
ATTTATGCTGATCAAATCTGCATTGGCTACCATGCCAACAACAGCACCGAGCAAGTTGAC
ACGATCATGGAGAAGAACGTAACCGTCACTCACGCTCAAGACATCCTGGAGAAGAAACAC
AACGGTAAGCTTTGCGACTTGGATGGAGTCAAGCCCCTCATACTTAGAGATTGTAGTGTA
GCCGGTTGGCTGCTCGGTAACCCAATGTGCGATGAGTTCATCAATGTTCCCGAATGGTCA
TATATCGTCGAAAAAGCTAATCCTGTCAACGACCTGTGCTACCCCGGTGATTTCAATGAC
TATGAAGAACTGAAGCACCTGCTCTCCCGCATCAACCATTTCGAGAAAATCCAGATCATT
CCCAAGAGTTCCTGGTCTAGCCATGAGGCTAGTCTGGGTGTCTCATCCGCCTGCCCATAT
CAGGGTAAAAGTTCTTTCTTTAGGAACGTAGTATGGTTGATAAAGAAAAACTCTACATAC
CCGACCATCAAGCGCTCTTACAACAATACGAACCAAGAGGATCTGCTTGTCCTTTGGGGA
ATCCATCATCCTAATGATGCTGCCGAACAGACTAAGCTCTACCAAAACCCTACCACTTAT
ATTTCCGTCGGCACCTCTACTCTGAACCAGCGCCTTGTGCCCAGGATCGCTACGAGATCA
AAAGTCAACGGCCAATCGGGCCGCATGGAATTCTTCTGGACGATCCTGAAGCCTAATGAC
GCTATCAACTTCGAGTCAAATGGAAACTTTATCGCTCCCGAGTACGCTTACAAGATCGTC
AAGAAGGGCGACTCCACGATTATGAAGTCAGAGTTGGAGTACGGCAACTGCAACACAAAG
TGCCAAACTCCTATGGGCGCTATAAATTCTTCAATGCCGTTCCACAACATCCATCCGCTC
ACGATCGGTGAGTGCCCGAAATATGTAAAGTCGAATCGTCTCGTACTGGCGACAGGCCTG
AGAAATAGTCCGCAACGTGAACGTCGTCGCAAGAAGAGAGGACTGTTTGGTGCCATTGCA
GGCTTTATTGAGGGCGGCTGGCAGGGCATGGTTGACGGATGGTACGGCTACCACCATTCA
AACGAGCAGGGATCTGGCTACGCCGCTGACAAAGAAAGCACCCAAAAGGCCATTGATGGA
GTGACGAATAAGGTGAATTCGATCATCGACAAAATGAACACGCAATTCGAAGCAGTGGGT
CGCGAATTCAATAACCTGGAGCGCCGTATCGAGAATCTGAACAAGAAGATGGAAGACGGC
TTTTTGGATGTCTGGACATATAACGCTGAATTGCTGGTCCTCATGGAAAACGAGCGTACC
CTTGATTTCCACGACAGCAACGTTAAGAACCTCTACGACAAGGTCAGGCTCCAGCTCAGG
GATAACGCCAAGGAATTGGGAAACGGATGCTTCGAGTTCTACCACAAATGCGACAACGAG
TGCATGGAGTCAGTCAGGAATGGTACCTACGACTACCCGCAATATTCTGAGGAGGCTCGC
TTGAAGCGTGAGGAAATATCGGGTGTTAAATTGGAGAGTATTGGAATCTACCAGCACCAT
CACCACCATCACTAATAGGCATGC

SEQ ID NO: 147 HA1-1 IND
ATGAAATTCTTAGTCAACGTTGCCCTTGTTTTTATGGTCGTATACATTTCTTACATCTAT
GCGGAGAAAACCCATAACGGTAAGTTGTGCGACCTTGACGGTGTAAAGCCCCTGATCCTC
CGTGACTGCAGTGTTGCTGGTTGGCTTTTGGGCAACCCCATGTGTGACGAATTTATCAAC
GTGCCTGAATGGTCATACATTGTAGAGAAGGCCAACCCCACGAACGATCTCTGTTATCCC

GGCAGCTTCAATGACTATGAGGAACTTAAGCACCTTCTGTCACGTATCAACCACTTCGAA
AAGATCCAGATCATCCCGAAGAGCTCCTGGAGCGACCACGAAGCCAGTTCGGGTGTGTCT
TCCGCTTGCCCCTACCTCGGTAGCCCTTCCTTCTTCCGTAACGTAGTGTGGCTGATCAAG
AAGAATAGCACTTACCCTACAATCAAAAAGTCGTATAACAATACTAACCAAGAGGATCTG
CTTGTACTCTGGGGAATTCATCATCCCAACGACGCGGCGGAGCAGACCAGGTTGTACCAG
AACCCCACCACTTACATCTCCATCGGTACGTCCACACTGAATCAGCGTCTGGTCCCCAAG
ATCGCAACCAGGTCCAAGGTTAACGGTCAGTCCGGTCGTATGGAGTTCTTCTGGACCATC
CTGAAGCCCAACGACGCCATCAACTTCGAGTCCAACGGTAATTTCATTGCTCCGGAGTAC
GCCTACAAGATAGTTAAGAAGGGTGATTCAGCGATCATGAAGTCGGAACTTGAGTATGGC
AACTGCAACACTAAATGCCAAACTCCAATGGGCGCTATCAACTCCAGTATGCCATTCCAT
AACATCCACCCATTGACAATCGGTGAATGTCCCAAGTACGTGAAGCACCACCATCACCAT
CACTAATAG

SEQ ID NO: 148 FOR PRIMER

AGGCAGATCTATGAAATTCTTAGTCAACGTTGCCCTTGTTTTTATGGTCGTATACATTTC
TTACATCTATGCGGAGAAAACCCATAACGGTAAGTTGTG

SEQ ID NO: 149 REV PRIMER

AGTCGCATGCCTATTAATGGTGGTGATGATGGTGCTTCACGTACTTGGGACATTCACCGA
TTG

SEQ ID NO: 150 HA0s IND

AGATCTATGAAGTTCCTGGTCAATGTAGCCTTGGTATTTATGGTAGTCTATATCTCGTAC
ATTTACGCAGACCAGATTTGTATTGGATATCACGCTAACAACAGCACAGAGCAGGTAGAT
ACTATTATGGAGAAAAATGTTACCGTCACTCACGCCCAGGACATCCTGGAGAAAACCCAT
AACGGTAAGTTGTGCGACCTTGACGGTGTAAAGCCCCTGATCCTCCGTGACTGCAGTGTT
GCTGGTTGGCTTTTGGGCAACCCCATGTGTGACGAATTTATCAACGTGCCTGAATGGTCA
TACATTGTAGAGAAGGCCAACCCCACGAACGATCTCTGTTATCCCGGCAGCTTCAATGAC
TATGAGGAACTTAAGCACCTTCTGTCACGTATCAACCACTTCGAAAAGATCCAGATCATC
CCGAAGAGCTCCTGGAGCGACCACGAAGCCAGTTCGGGTGTGTCTTCCGCTTGCCCCTAC
CTCGGTAGCCCTTCCTTCTTCCGTAACGTAGTGTGGCTGATCAAGAAGAATAGCACTTAC
CCTACAATCAAAAAGTCGTATAACAATACTAACCAAGAGGATCTGCTTGTACTCTGGGGA
ATTCATCATCCCAACGACGCGGCGGAGCAGACCAGGTTGTACCAGAACCCCACCACTTAC
ATCTCCATCGGTACGTCCACACTGAATCAGCGTCTGGTCCCCAAGATCGCAACCAGGTCC
AAGGTTAACGGTCAGTCCGGTCGTATGGAGTTCTTCTGGACCATCCTGAAGCCCAACGAC
GCCATCAACTTCGAGTCCAACGGTAATTTCATTGCTCCGGAGTACGCCTACAAGATAGTT
AAGAAGGGTGATTCAGCGATCATGAAGTCGGAACTTGAGTATGGCAACTGCAACACTAAA
TGCCAAACTCCAATGGGCGCTATCAACTCCAGTATGCCATTCCATAACATCCACCCATTG
ACAATCGGTGAATGTCCCAAGTACGTGAAGAGCAACAGGTTGGTATTGGCCACCGGTTTG
AGAAACAGCCCCCAGAGAGAGTCGCGTCGTAAAAAGCGCGGCTTGTTCGGAGCCATCGCT
GGCTTCATAGAGGGTGGTTGGCAGGGAATGGTCGATGGTTGGTATGGTTATCATCATTCC
AACGAGCAGGGAAGTGGTTACGCCGCCGACAAAGAATCGACCCAGAAGGCTATTGACGGC
GTCACAAACAAAGTAAACTCTATCATTGATAAGATGAACACCCAGTTCGAGGCTGTAGGT
AGAGAATTCAACAACCTCGAAAGACGTATTGAGAACCTGAACAAGAAAATGGAGGATGGC
TTCCTGGACGTGTGGACCTACAATGCTGAGCTGTTGGTCCTTATGGAGAACGAGCGTACC
CTCGATTTCCATGACTCAAACGTGAAGAACCTGTATGACAAGGTGCGTTTGCAACTGAGG
GACAACGCAAAGGAGCTTGGAAACGGTTGTTTCGAATTTTATCATAAGTGCGACAATGAG
TGTATGGAGTCGATTAGAAATGGCACGTACAACTACCCTCAATACAGCGAAGAAGCTCGT
CTCAAACGTGAGGAAATCAGCGGCGTCAAGCTCGAATCAATCGGTACCTATCAGCACCAC
CATCACCATCACTAATAGGCATGC

SEQ ID NO: 151 STF2.HA1-1His(PR8)

MKFLVNVALVFMVVYISYIYAAQVINTNSLSLLTQNNLNKSQSALGTAIERLSSGLRINSA
KDDAAGQAIANRFTANIKGLTQASRNANDGISIAQTTEGALNEINNNLQRVRELAVQSANS
TNSQSDLDSIQAEITQRLNEIDRVSGQTQFNGVKVLAQDNTLTIQVGANDGETIDIDLKQI
NSQTLGLDSLNVQKAYDVKDTAVTTKAYANNGTTLDVSGLDDAAIKAATGGTNGTASVTGG
AVKFDADNNKYFVTIGGFTGADAAKNGDYEVNVATDGTVTLAAGATKTTMPAGATTKTEVQ
ELKDTPAVVSADAKNALIAGGVDATDANGAELVKMSYTDKNGKTIEGGYALKAGDKYYAAD
YDEATGAIKAKTTSYTAADGTTKTAANQLGGVDGKTEVVTIDGKTYNASKAAGHDFKAQPE
LAEAAAKTTENPLQKIDAALAQVDALRSDLGAVQNRFNSAITNLGNTVNNLSEARSRIEDS
DYATEVSNMSRAQILQQAGTSVLAQANQVPQNVLSLLASGSGSGSSHNGKLCRLKGIAPLQ
LGKCNIAGWLLGNPECDPLLPVRSWSYIVETPNSENGICYPGDFIDYEELREQLSSVSSFE
RFEIFPKESSWPNHNTNGVTAACSHEGKSSFYRNLLWLTEKEGSYPKLKNSYVNKKGKEVL
VLWGIHHPPNSKEQQNLYQNENAYVSVVTSNYNRRFTPEIAERPKVRDQAGRMNYYWTLLK
PGDTIIFEANGNLIAPMYAFALSRGFGSGIITSNASMHECNTKCQTPLGAINSSLPYQNIH
PVTIGECPKYVRHHHHHH

SEQ ID NO: 152 ngSTF2.HA1-1His(PR8)

MKFLVNVALVFMVVYISYIYAAQVINTNSLSLLTQNNLQKSQSALGTAIERLSSGLRINSA
KDDAAGQAIANRFTANIKGLTQASRNANDGISIAQTTEGALNEINNNLQRVRELAVQSAQS
TNSQSDLDSIQAEITQRLNEIDRVSGQTQFNGVKVLAQDNTLTIQVGANDGETIDIDLKQI
NSQTLGLDSLNVQKAYDVKDTAVTTKAYANQGTTLDVSGLDDAAIKAATGGTQGTASVTGG
AVKFDADNNKYFVTIGGFTGADAAKNGDYEVNVATDGTVTLAAGATKTTMPAGATTKTEVQ
ELKDTPAVVSADAKNALIAGGVDATDANGAELVKMSYTDKNGKTIEGGYALKAGDKYYAAD
YDEATGAIKAKTTSYTAADGTTKTAANQLGGVDGKTEVVTIDGKTYQASKAAGHDFKAQPE
LAEAAAKTTENPLQKIDAALAQVDALRSDLGAVQNRFNSAITNLGNTVNQLSEARSRIEDS
DYATEVSQMSRAQILQQAGTSVLAQANQVPQNVLSLLASGSGSGSSHNGKLCRLKGIAPLQ
LGKCNIAGWLLGNPECDPLLPVRSWSYIVETPNSENGICYPGDFIDYEELREQLSSVSSFE
RFEIFPKESSWPNHNTNGVTAACSHEGKSSFYRNLLWLTEKEGSYPKLKNSYVNKKGKEVL
VLWGIHHPPNSKEQQNLYQNENAYVSVVTSNYNRRFTPEIAERPKVRDQAGRMNYYWTLLK
PGDTIIFEANGNLIAPMYAFALSRGFGSGIITSQASMHECNTKCQTPLGAIQSSLPYQNIH
PVTIGECPKYVRHHHHHH

SEQ ID NO: 153 STF2.HA1-2His(PR8)

MKFLVNVALVFMVVYISYIYAAQVINTNSLSLLTQNNLNKSQSALGTAIERLSSGLRINSA
KDDAAGQAIANRFTANIKGLTQASRNANDGISIAQTTEGALNEINNNLQRVRELAVQSANS
TNSQSDLDSIQAEITQRLNEIDRVSGQTQFNGVKVLAQDNTLTIQVGANDGETIDIDLKQI
NSQTLGLDSLNVQKAYDVKDTAVTTKAYANNGTTLDVSGLDDAAIKAATGGTNGTASVTGG
AVKFDADNNKYFVTIGGFTGADAAKNGDYEVNVATDGTVTLAAGATKTTMPAGATTKTEVQ
ELKDTPAVVSADAKNALIAGGVDATDANGAELVKMSYTDKNGKTIEGGYALKAGDKYYAAD
YDEATGAIKAKTTSYTAADGTTKTAANQLGGVDGKTEVVTIDGKTYNASKAAGHDFKAQPE
LAEAAAKTTENPLQKIDAALAQVDALRSDLGAVQNRFNSAITNLGNTVNNLSEARSRIEDS
DYATEVSNMSRAQILQQAGTSVLAQANQVPQNVLSLLASGSGSGSKGIAPQLGKCNIAGWL
LGNPECDPLLPVRSWSYIVETPNSENGICYPGDFIDYEELREQLSSVSSFERFEIFPKESS
WPNHNTNGVTAACSHEGKSSFYRNLLWLTEKEGSYPKLKNSYVNKKGKEVLVLWGIHHPPN
SKEQQNLYQNENAYVSVVTSNYNRRFTPEIAERPKVRDQAGRMNYYWTLLKPGDTIIFEAN
GNLIAPMYAFALSRGFGSGIITSHHHHHH

SEQ ID NO: 154 ngSTF2.HA1-2His(PR8)

MKFLVNVALVFMVVYISYIYAAQVINTNSLSLLTQNNLQKSQSALGTAIERLSSGLRINSA
KDDAAGQAIANRFTANIKGLTQASRNANDGISIAQTTEGALNEINNNLQRVRELAVQSAQS
TNSQSDLDSIQAEITQRLNEIDRVSGQTQFNGVKVLAQDNTLTIQVGANDGETIDIDLKQI
NSQTLGLDSLNVQKAYDVKDTAVTTKAYANQGTTLDVSGLDDAAIKAATGGTQGTASVTGG
AVKFDADNNKYFVTIGGFTGADAAKNGDYEVNVATDGTVTLAAGATKTTMPAGATTKTEVQ

ELKDTPAVVSADAKNALIAGGVDATDANGAELVKMSYTDKNGKTIEGGYALKAGDKYYAAD
YDEATGAIKAKTTSYTAADGTTKTAANQLGGVDGKTEVVTIDGKTYQASKAAGHDFKAQPE
LAEAAAKTTENPLQKIDAALAQVDALRSDLGAVQNRFNSAITNLGNTVNQLSEARSRIEDS
DYATEVSQMSRAQILQQAGTSVLAQANQVPQNVLSLLASGSGSGSKGIAPLQLGKCNIAGW
LLGNPECDPLLPVRSWSYIVETPNSENGICYPGDFIDYEELREQLSSVSSFERFEIFPKES
SWPNHNTNGVTAACSHEGKSSFYRNLLWLTEKEGSYPKLKNSYVNKKGKEVLVLWGIHHPP
NSKEQQNLYQNENAYVSVVTSNYNRRFTPEIAERPKVRDQAGRMNYYWTLLKPGDTIIFEA
NGNLIAPMYAFALSRGFGSGIITSHHHHHH

SEQ ID NO: 155 STF2.HA1-2*mut*His(PR8)

MKFLVNVALVFMVVYISYIYAAQVINTNSLSLLTQNNLNKSQSALGTAIERLSSGLRINSA
KDDAAGQAIANRFTANIKGLTQASRNANDGISIAQTTEGALNEINNNLQRVRELAVQSANS
TNSQSDLDSIQAEITQRLNEIDRVSGQTQFNGVKVLAQDNTLTIQVGANDGETIDIDLKQI
NSQTLGLDSLNVQKAYDVKDTAVTTKAYANNGTTLDVSGLDDAAIKAATGGTNGTASVTGG
AVKFDADNNKYFVTIGGFTGADAAKNGDYEVNVATDGTVTLAAGATKTTMPAGATTKTEVQ
ELKDTPAVVSADAKNALIAGGVDATDANGAELVKMSYTDKNGKTIEGGYALKAGDKYYAAD
YDEATGAIKAKTTSYTAADGTTKTAANQLGGVDGKTEVVTIDGKTYNASKAAGHDFKAQPE
LAEAAAKTTENPLQKIDAALAQVDALRSDLGAVQNRFNSAITNLGNTVNNLSEARSRIEDS
DYATEVSNMSRAQILQQAGTSVLAQANQVPQNVLSLLASGSGSGSKGAAPLQLGKCNIAGW
LLGNPECDPLLPVRSWSDIAETPNSENGICYPGDFIDYEELREQLSSVSSFERFEIFPKES
SWPNHNTNGVTAACSHEGKSSFYRNLLWLTEKEGSYPKLKNSYVNKKGKEVLVLWGIHHPP
NSKEQQNLYQNENAYVSVVTSNYNRRFTPEIAERPKVRDQAGRMNYYWTLLKPGDTIIFEA
NGNLIAPMYAFALSRGFGSGIITSHHHHHH

SEQ ID NO: 156 ng*STF2.HA1-2mut*His(PR8)

MKFLVNVALVFMVVYISYIYAAQVINTNSLSLLTQNNLQKSQSALGTAIERLSSGLRINSA
KDDAAGQAIANRFTANIKGLTQASRNANDGISIAQTTEGALNEINNNLQRVRELAVQSAQS
TNSQSDLDSIQAEITQRLNEIDRVSGQTQFNGVKVLAQDNTLTIQVGANDGETIDIDLKQI
NSQTLGLDSLNVQKAYDVKDTAVTTKAYANQGTTLDVSGLDDAAIKAATGGTQGTASVTGG
AVKFDADNNKYFVTIGGFTGADAAKNGDYEVNVATDGTVTLAAGATKTTMPAGATTKTEVQ
ELKDTPAVVSADAKNALIAGGVDATDANGAELVKMSYTDKNGKTIEGGYALKAGDKYYAAD
YDEATGAIKAKTTSYTAADGTTKTAANQLGGVDGKTEVVTIDGKTYQASKAAGHDFKAQPE
LAEAAAKTTENPLQKIDAALAQVDALRSDLGAVQNRFNSAITNLGNTVNQLSEARSRIEDS
DYATEVSQMSRAQILQQAGTSVLAQANQVPQNVLSLLASGSGSGSKGAAPQLGKCNIAGWL
LGNPECDPLLPVRSWSDIAETPNSENGICYPGDFIDYEELREQLSSVSSFERFEIFPKESS
WPNHNTNGVTAACSHEGKSSFYRNLLWLTEKEGSYPKLKNSYVNKKGKEVLVLWGIHHPPN
SKEQQNLYQNENAYVSVVTSNYNRRFTPEIAERPKVRDQAGRMNYYWTLLKPGDTIIFEAN
GNLIAPMYAFALSRGFGSGIITSHHHHHH

SEQ ID NO: 157 STF2.HA1-3His(PR8)

MKFLVNVALVFMVVYISYIYAAQVINTNSLSLLTQNNLNKSQSALGTAIERLSSGLRINSA
KDDAAGQAIANRFTANIKGLTQASRNANDGISIAQTTEGALNEINNNLQRVRELAVQSANS
TNSQSDLDSIQAEITQRLNEIDRVSGQTQFNGVKVLAQDNTLTIQVGANDGETIDIDLKQI
NSQTLGLDSLNVQKAYDVKDTAVTTKAYANNGTTLDVSGLDDAAIKAATGGTNGTASVTGG
AVKFDADNNKYFVTIGGFTGADAAKNGDYEVNVATDGTVTLAAGATKTTMPAGATTKTEVQ
ELKDTPAVVSADAKNALIAGGVDATDANGAELVKMSYTDKNGKTIEGGYALKAGDKYYAAD
YDEATGAIKAKTTSYTAADGTTKTAANQLGGVDGKTEVVTIDGKTYNASKAAGHDFKAQPE
LAEAAAKTTENPLQKIDAALAQVDALRSDLGAVQNRFNSAITNLGNTVNNLSEARSRIEDS
DYATEVSNMSRAQILQQAGTSVLAQANQVPQNVLSLLASGSGSGSNSENGCYPGDFIDYEE
LREQLSSVSSFERFEIFPKESSWPNHNTNGVTAACSHEGKSSFYRNLLWLTEKEGSYPKLK
NSYVNKKGKEVLVLWGIHHPPNSKEQQNLYQNENAYVSVVTSNYNRRFTPEIAERPKVRDQ
AGRMNYYWTLLKPGDTIIFEANGNLIAPMYAFALSRGHHHHHH

SEQ ID NO: 158 STF2.HA1-3*mut*His(PR8)

MKFLVNVALVFMVVYISYIYAAQVINTNSLSLLTQNNLNKSQSALGTAIERLSSGLRINSA
KDDAAGQAIANRFTANIKGLTQASRNANDGISIAQTTEGALNEINNNLQRVRELAVQSANS
TNSQSDLDSIQAEITQRLNEIDRVSGQTQFNGVKVLAQDNTLTIQVGANDGETIDIDLKQI
NSQTLGLDSLNVQKAYDVKDTAVTTKAYANNGTTLDVSGLDDAAIKAATGGTNGTASVTGG
AVKFDADNNKYFVTIGGFTGADAAKNGDYEVNVATDGTVTLAAGATKTTMPAGATTKTEVQ
ELKDTPAVVSADAKNALIAGGVDATDANGAELVKMSYTDKNGKTIEGGYALKAGDKYYAAD
YDEATGAIKAKTTSYTAADGTTKTAANQLGGVDGKTEVVTIDGKTYNASKAAGHDFKAQPE
LAEAAAKTTENPLQKIDAALAQVDALRSDLGAVQNRFNSAITNLGNTVNNLSEARSRIEDS
DYATEVSNMSRAQILQQAGTSVLAQANQVPQNVLSLLASGSGSGSNSENEICYPGDFIDKE
ELREQLSSVSSFERFEIFPKESSWPNHNTNGVTAACSHEGKSSFYRNLLWLTEKEGSYPKL
KNSYVNKKGKEVLVLWGIHHPPNSKEQQNLYQNENAYVSVVTSNYNRRFTPEIAERPKVRD
QAGRMNYYWTLLKPGDTIIFEANGNLIAPMYAAALSRGHHHHHH

SEQ ID NO: 159 STF2.HA1-2(IND)

MAQVINTNSLSLLTQNNLNKSQSALGTAIERLSSGLRINSAKDDAAGQAIANRFTANIKGL
TQASRNANDGISIAQTTEGALNEINNNLQRVRELAVQSANSTNSQSDLDSIQAEITQRLNE
IDRVSGQTQFNGVKVLAQDNTLTIQVGANDGETIDIDLKQINSQTLGLDSLNVQKAYDVKD
TAVTTKAYANNGTTLDVSGLDDAAIKAATGGTNGTASVTGGAVKFDADNNKYFVTIGGFTG
ADAAKNGDYEVNVATDGTVTLAAGATKTTMPAGATTKTEVQELKDTPAVVSADAKNALIAG
GVDATDANGAELVKMSYTDKNGKTIEGGYALKAGDKYYAADYDEATGAIKAKTTSYTAADG
TTKTAANQLGGVDGKTEVVTIDGKTYNASKAAGHDFKAQPELAEAAAKTTENPLQKIDAAL
AQVDALRSDLGAVQNRFNSAITNLGNTVNNLSEARSRIEDSDYATEVSNMSRAQILQQAGT
SVLAQANQVPQNVLSLLAGVKPLILRDCSVAGWLLGNPMCDEFINVPEWSYIVEKANPTND
LCYPGSFNDYEELKHLLSRINHFEKIQIIPKSSWSDHEASSGVSSACPYLGSPSFFRNVVW
LIKKNSTYPTIKKSYNNTNQEDLLVLWGIHHPNDAAEQTRLYQNPTTYISIGTSTLNQRLV
PKIATRSKVNGQSGRMEFFWTILKPNDAINFESNGNFIAPEYAYKIVKKGDSAIMKSE

Drosophila

SEQ ID NO: 160 STF2Δ.HA0shis

ATGGCACAAGTAATCAACACTAACAGTCTGTCGCTGCTGACCCAGAATAACCTGAACAAA
TCCCAGTCCGCACTGGGCACCGCTATCGAGCGTCTGTCTTCTGGTCTGCGTATCAACAGC
GCGAAAGACGATGCGGCAGGTCAGGCGATTGCTAACCGTTTCACCGCGAACATCAAAGGT
CTGACTCAGGCTTCCCGTAACGCTAACGACGGTATCTCCATTGCGCAGACCACTGAAGGC
GCGCTGAACGAAATCAACAACAACCTGCAGCGTGTGCGTGAACTGGCGGTTCAGTCTGCT
AACAGCACCAACTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATCACCCAGCGCCTG
AACGAAATCGACCGTGTATCCGGCCAGACTCAGTTCAACGGCGTGAAAGTCCTGGCGCAG
GACAACACCCTGACCATCCAGGTTGGCGCCAACGACGGTGAAACTATCGATATCGATCTG
AAGCAGATCAACTCTCAGACCCTGGGTCTGGACTCACTGAACGTGCATGGAGCGCCGGTG
GATCCTGCTAGCCCATGGACCGAAAACCCGCTGCAGAAAATTGATGCCGCGCTGGCGCAG
GTGGATGCGCTGCGCTCTGATCTGGGTGCGGTACAAAACCGTTTCAACTCTGCTATCACC
AACCTGGGCAATACCGTAAACAATCTGTCTGAAGCGCGTAGCCGTATCGAAGATTCCGAC
TACGCGACCGAAGTTTCCAACATGTCTCGCGCGCAGATTTTGCAGCAGGCCGGTACTTCC
GTTCTGGCGCAGGCTAACCAGGTCCCGCAGAACGTGCTGTCTCTGTTACGTGAATTCTCT
AGATATCCAGCACAGTGGCGGCCGCTCGACACAATATGTATAGGCTACCATGCGAACAAT
TCAACCGACACTGTTGACACAGTACTCGAGAAGAATGTGACAGTGACACACTCTGTTAAC
CTGCTCGAAGACAGCCACAACGGAAAACTATGTAGATTAAAAGGAATAGCCCCACTACAA
TTGGGGAAATGTAACATCGCCGGATGGCTTTTGGGAAACCCAGAATGCGACCCACTGCTT
CCAGTGAGATCATGGTCCTACATTGTAGAAACACCAAACTCTGAGAATGGAATATGTTAT
CCAGGAGATTTCATCGACTATGAGGAGCTGAGGGAGCAATTGAGCTCAGTGTCATCATTC
GAAAGATTCGAAATATTTCCCAAAGAAAGCTCATGGCCCAACCACAACACAAACGGAGTA
ACGGCAGCATGCTCCCATGAGGGGAAAAGCAGTTTTTACAGAAATTTGCTATGGCTGACG
GAGAAGGAGGGCTCATACCCAAAGCTGAAAAATTCTTATGTGAACAAAAAAGGGAAAGAA

GTCCTTGTACTGTGGGGTATTCATCACCCGCCTAACAGTAAGGAACAACAGAATCTCTAT
CAGAATGAAAATGCTTATGTCTCTGTAGTGACTTCAAATTATAACAGGAGATTTACCCCG
GAAATAGCAGAAAGACCCAAAGTAAGAGATCAAGCTGGGAGGATGAACTATTACTGGACC
TTGCTAAAACCCGGAGACACAATAATATTTGAGGCAAATGGAAATCTAATAGCACCAATG
TATGCTTTCGCACTGAGTAGAGGCTTTGGGTCCGGCATCATCACCTCAAACGCATCAATG
CATGAGTGTAACACGAAGTGTCAAACACCCCTGGGAGCTATAAACAGCAGTCTCCCTTAC
CAGAATATACACCCAGTCACAATAGGAGAGTGCCCAAAATACGTCAGGAGTGCCAAATTG
AGGATGGTTACAGGACTAAGGAACATTCCGTCCATTCAATCCAGAGGTCTATTTGGAGCC
ATTGCCGGTTTTATTGAAGGGGGATGGACTGGAATGATAGATGGATGGTATGGTTATCAT
CATCAGAATGAACAGGGATCAGGCTATGCAGCGGATCAAAAAAGCACACAAAATGCCATT
AACGGGATTACAAACAAGGTGAACACTGTTATCGAGAAAATGAACATTCAATTCACAGCT
GTGGGTAAAGAATTCAACAAATTAGAAAAAAGGATGGAAAATTTAAATAAAAAAGTTGAT
GATGGATTTCTGGACATTTGGACATATAATGCAGAATTGTTAGTTCTACTGGAAAATGAA
AGGACTCTGGATTTCCATGACTCAAATGTGAAGAATCTGTATGAGAAAGTAAAAAGCCAA
TTAAAGAATAATGCCAAAGAAATCGGAAATGGATGTTTTGAGTTCTACCACAAGTGTGAC
AATGAATGCATGGAAAGTGTAAGAAATGGGACTTATGATTATCCCAAATATTCAGAAGAG
TCAAAGTTGAACAGGGAAAAGGTAGATGGAGTGAAATTGGAATCAATGGGGATCTATCAG
ACGCGTACCGGTCATCATCACCATCACCATTGA

SEQ ID NO: 161 FOR PRIMER
ATTCTCCGGCGGCCGCTCGACACAATATGTATAGGCTACC

SEQ ID NO: 162 REV PRIMER
AGTCTTGCGGCCGCCTATTAATGGTGATGGTGATGATGCTGATAGATCCCCATTGATTCC

SEQ ID NO: 163 HA0s PR8 template
GACACAATATGTATAGGCTACCATGCGAACAATTCAACCGACACTGTTGACACAGTACTC
GAGAAGAATGTGACAGTGACACACTCTGTTAACCTGCTCGAAGACAGCCACAACGGAAAA
CTATGTAGATTAAAAGGAATAGCCCCACTACAATTGGGGAAATGTAACATCGCCGGATGG
CTTTTGGGAAACCCAGAATGCGACCCACTGCTTCCAGTGAGATCATGGTCCTACATTGTA
GAAACACCAAACTCTGAGAATGGAATATGTTATCCAGGAGATTTCATCGACTATGAGGAG
CTGAGGGAGCAATTGAGCTCAGTGTCATCATTCGAAAGATTCGAAATATTTCCCAAAGAA
AGCTCATGGCCCAACCACAACACAAACGGAGTAACGGCAGCATGCTCCCATGAGGGGAAA
AGCAGTTTTTACAGAAATTTGCTATGGCTGACGGAGAAGGAGGGCTCATACCCAAAGCTG
AAAAATTCTTATGTGAACAAAAAAGGGAAAGAAGTCCTTGTACTGTGGGGTATTCATCAC
CCGCCTAACAGTAAGGAACAACAGAATCTCTATCAGAATGAAAATGCTTATGTCTCTGTA
GTGACTTCAAATTATAACAGGAGATTTACCCCGGAAATAGCAGAAAGACCCAAAGTAAGA
GATCAAGCTGGGAGGATGAACTATTACTGGACCTTGCTAAAACCCGGAGACACAATAATA
TTTGAGGCAAATGGAAATCTAATAGCACCAATGTATGCTTTCGCACTGAGTAGAGGCTTT
GGGTCCGGCATCATCACCTCAAACGCATCAATGCATGAGTGTAACACGAAGTGTCAAACA
CCCCTGGGAGCTATAAACAGCAGTCTCCCTTACCAGAATATACACCCAGTCACAATAGGA
GAGTGCCCAAAATACGTCAGGAGTGCCAAATTGAGGATGGTTACAGGACTAAGGAACATT
CCGTCCATTCAATCCAGAGGTCTATTTGGAGCCATTGCCGGTTTTATTGAAGGGGGATGG
ACTGGAATGATAGATGGATGGTATGGTTATCATCATCAGAATGAACAGGGATCAGGCTAT
GCAGCGGATCAAAAAAGCACACAAAATGCCATTAACGGGATTACAAACAAGGTGAACACT
GTTATCGAGAAAATGAACATTCAATTCACAGCTGTGGGTAAAGAATTCAACAAATTAGAA
AAAAGGATGGAAAATTTAAATAAAAAAGTTGATGATGGATTTCTGGACATTTGGACATAT
AATGCAGAATTGTTAGTTCTACTGGAAAATGAAAGGACTCTGGATTTCCATGACTCAAAT
GTGAAGAATCTGTATGAGAAAGTAAAAAGCCAATTAAAGAATAATGCCAAAGAAATCGGA
AATGGATGTTTTGAGTTCTACCACAAGTGTGACAATGAATGCATGGAAAGTGTAAGAAAT
GGGACTTATGATTATCCCAAATATTCAGAAGAGTCAAAGTTGAACAGGGAAAAGGTAGAT
GGAGTGAAATTGGAATCAATGGGGATCTATCAG

SEQ ID NO: 164 STF2Δ.HA0s

ATGGCACAAGTAATCAACACTAACAGTCTGTCGCTGCTGACCCAGAATAACCTGAACAAA
TCCCAGTCCGCACTGGGCACCGCTATCGAGCGTCTGTCTTCTGGTCTGCGTATCAACAGC
GCGAAAGACGATGCGGCAGGTCAGGCGATTGCTAACCGTTTCACCGCGAACATCAAAGGT
CTGACTCAGGCTTCCCGTAACGCTAACGACGGTATCTCCATTGCGCAGACCACTGAAGGC
GCGCTGAACGAAATCAACAACAACCTGCAGCGTGTGCGTGAACTGGCGGTTCAGTCTGCT
AACAGCACCAACTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATCACCCAGCGCCTG
AACGAAATCGACCGTGTATCCGGCCAGACTCAGTTCAACGGCGTGAAAGTCCTGGCGCAG
GACAACACCCTGACCATCCAGGTTGGCGCCAACGACGGTGAAACTATCGATATCGATCTG
AAGCAGATCAACTCTCAGACCCTGGGTCTGGACTCACTGAACGTGCATGGAGCGCCGGTG
GATCCTGCTAGCCCATGGACCGAAAACCCGCTGCAGAAAATTGATGCCGCGCTGGCGCAG
GTGGATGCGCTGCGCTCTGATCTGGGTGCGGTACAAAACCGTTTCAACTCTGCTATCACC
AACCTGGGCAATACCGTAAACAATCTGTCTGAAGCGCGTAGCCGTATCGAAGATTCCGAC
TACGCGACCGAAGTTTCCAACATGTCTCGCGCGCAGATTTTGCAGCAGGCCGGTACTTCC
GTTCTGGCGCAGGCTAACCAGGTCCCGCAGAACGTGCTGTCTCTGTTACGTGAATTCTCT
AGATATCCAGCACAGTGGCGGCCGCTCGACACAATATGTATAGGCTACCATGCGAACAAT
TCAACCGACACTGTTGACACAGTACTCGAGAAGAATGTGACAGTGACACACTCTGTTAAC
CTGCTCGAAGACAGCCACAACGGAAAACTATGTAGATTAAAAGGAATAGCCCCACTACAA
TTGGGGAAATGTAACATCGCCGGATGGCTTTTGGGAAACCCAGAATGCGACCCACTGCTT
CCAGTGAGATCATGGTCCTACATTGTAGAAACACCAAACTCTGAGAATGGAATATGTTAT
CCAGGAGATTTCATCGACTATGAGGAGCTGAGGGAGCAATTGAGCTCAGTGTCATCATTC
GAAAGATTCGAAATATTTCCCAAAGAAAGCTCATGGCCCAACCACAACACAAACGGAGTA
ACGGCAGCATGCTCCCATGAGGGGAAAAGCAGTTTTTACAGAAATTTGCTATGGCTGACG
GAGAAGGAGGGCTCATACCCAAAGCTGAAAAATTCTTATGTGAACAAAAAAGGGAAAGAA
GTCCTTGTACTGTGGGGTATTCATCACCCGCCTAACAGTAAGGAACAACAGAATCTCTAT
CAGAATGAAAATGCTTATGTCTCTGTAGTGACTTCAAATTATAACAGGAGATTTACCCCG
GAAATAGCAGAAAGACCCAAAGTAAGAGATCAAGCTGGGAGGATGAACTATTACTGGACC
TTGCTAAAACCCGGAGACACAATAATATTTGAGGCAAATGGAAATCTAATAGCACCAATG
TATGCTTTCGCACTGAGTAGAGGCTTTGGGTCCGGCATCATCACCTCAAACGCATCAATG
CATGAGTGTAACACGAAGTGTCAAACACCCCTGGGAGCTATAAACAGCAGTCTCCCTTAC
CAGAATATACACCCAGTCACAATAGGAGAGTGCCCAAAATACGTCAGGAGTGCCAAATTG
AGGATGGTTACAGGACTAAGGAACATTCCGTCCATTCAATCCAGAGGTCTATTTGGAGCC
ATTGCCGGTTTTATTGAAGGGGGATGGACTGGAATGATAGATGGATGGTATGGTTATCAT
CATCAGAATGAACAGGGATCAGGCTATGCAGCGGATCAAAAAAGCACACAAAATGCCATT
AACGGGATTACAAACAAGGTGAACACTGTTATCGAGAAAATGAACATTCAATTCACAGCT
GTGGGTAAAGAATTCAACAAATTAGAAAAAAGGATGGAAAATTTAAATAAAAAAGTTGAT
GATGGATTTCTGGACATTTGGACATATAATGCAGAATTGTTAGTTCTACTGGAAAATGAA
AGGACTCTGGATTTCCATGACTCAAATGTGAAGAATCTGTATGAGAAAGTAAAAAGCCAA
TTAAAGAATAATGCCAAAGAAATCGGAAATGGATGTTTTGAGTTCTACCACAAGTGTGAC
AATGAATGCATGGAAAGTGTAAGAAATGGGACTTATGATTATCCCAAATATTCAGAAGAG
TCAAAGTTGAACAGGGAAAAGGTAGATGGAGTGAAATTGGAATCAATGGGGATCTATCAG
TAG

SEQ ID NO: 165 FOR PRIMER

ATTCTCCGGCGGCCGCTCGACACAATATGTATAGGCTACC

SEQ ID NO: 166 REV PRIMER

AGTCTTGCGGCCGCCTATTAATGGTGATGGTGATGATGCTGATAGATCCCCATTGATTCC

SEQ ID NO: 167 HA0s.STF2Δ

GACACAATATGTATAGGCTACCATGCGAACAATTCAACCGACACTGTTGACACAGTACTC
GAGAAGAATGTGACAGTGACACACTCTGTTAACCTGCTCGAAGACAGCCACAACGGAAAA
CTATGTAGATTAAAAGGAATAGCCCCACTACAATTGGGGAAATGTAACATCGCCGGATGG

CTTTTGGGAAACCCAGAATGCGACCCACTGCTTCCAGTGAGATCATGGTCCTACATTGTA
GAAACACCAAACTCTGAGAATGGAATATGTTATCCAGGAGATTTCATCGACTATGAGGAG
CTGAGGGAGCAATTGAGCTCAGTGTCATCATTCGAAAGATTCGAAATATTTCCCAAAGAA
AGCTCATGGCCCAACCACAACACAAACGGAGTAACGGCAGCATGCTCCCATGAGGGGAAA
AGCAGTTTTTACAGAAATTTGCTATGGCTGACGGAGAAGGAGGGCTCATACCCAAAGCTG
AAAAATTCTTATGTGAACAAAAAAGGGAAAGAAGTCCTTGTACTGTGGGGTATTCATCAC
CCGCCTAACAGTAAGGAACAACAGAATCTCTATCAGAATGAAAATGCTTATGTCTCTGTA
GTGACTTCAAATTATAACAGGAGATTTACCCCGGAAATAGCAGAAAGACCCAAAGTAAGA
GATCAAGCTGGGAGGATGAACTATTACTGGACCTTGCTAAAACCCGGAGACACAATAATA
TTTGAGGCAAATGGAAATCTAATAGCACCAATGTATGCTTTCGCACTGAGTAGAGGCTTT
GGGTCCGGCATCATCACCTCAAACGCATCAATGCATGAGTGTAACACGAAGTGTCAAACA
CCCCTGGGAGCTATAAACAGCAGTCTCCCTTACCAGAATATACACCCAGTCACAATAGGA
GAGTGCCCAAAATACGTCAGGAGTGCCAAGTTGAGGATGGTTACAGGACTAAGGAACATT
CCGTCCATTCAATCCAGAGGTCTATTTGGAGCCATTGCCGGTTTTATTGAAGGGGGATGG
ACTGGAATGATAGATGGATGGTATGGTTATCATCATCAGAATGAACAGGGATCAGGCTAT
GCAGCGGATCAAAAAAGCACACAAAATGCCATTAACGGGATTACAAACAAGGTGAACACT
GTTATCGAGAAAATGAACATTCAATTCACAGCTGTGGGTAAAGAATTCAACAAATTAGAA
AAAAGGATGGAAAATTTAAATAAAAAAGTTGATGATGGATTTCTGGACATTTGGACATAT
AATGCAGAATTGTTAGTTCTACTGGAAAATGAAAGGACTCTGGATTTCCATGACTCAAAT
GTGAAGAATCTGTATGAGAAAGTAAAAAGCCAATTAAAGAATAATGCCAAAGAAATCGGA
AATGGATGTTTTGAGTTCTACCACAAGTGTGACAATGAATGCATGGAAAGTGTAAGAAAT
GGGACTTATGATTATCCCAAATATTCAGAAGAGTCAAAGTTGAACAGGGAAAAGGTAGAT
GGAGTGAAATTGGAATCAATGGGGATCTATCAGGAATTCTCTAGATATCCAGCACAGTGG
CGGCCGCTCGCACAAGTAATCAACACTAACAGTCTGTCGCTGCTGACCCAGAATAACCTG
AACAAATCCCAGTCCGCACTGGGCACCGCTATCGAGCGTCTGTCTTCTGGTCTGCGTATC
AACAGCGCGAAAGACGATGCGGCAGGTCAGGCGATTGCTAACCGTTTCACCGCGAACATC
AAAGGTCTGACTCAGGCTTCCCGTAACGCTAACGACGGTATCTCCATTGCGCAGACCACT
GAAGGCGCGCTGAACGAAATCAACAACAACCTGCAGCGTGTGCGTGAACTGGCGGTTCAG
TCTGCTAACAGCACCAACTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATCACCCAG
CGCCTGAACGAAATCGACCGTGTATCCGGCCAGACTCAGTTCAACGGCGTGAAAGTCCTG
GCGCAGGACAACACCCTGACCATCCAGGTTGGCGCCAACGACGGTGAAACTATCGATATC
GATCTGAAGCAGATCAACTCTCAGACCCTGGGTCTGGACTCACTGAACGTGCATGGAGCG
CCGGTGGATCCTGCTAGCCCATGGACCGAAAACCCGCTGCAGAAAATTGATGCCGCGCTG
GCGCAGGTGGATGCGCTGCGCTCTGATCTGGGTGCGGTACAAAACCGTTTCAACTCTGCT
ATCACCAACCTGGGCAATACCGTAAACAATCTGTCTGAAGCGCGTAGCCGTATCGAAGAT
TCCGACTACGCGACCGAAGTTTCCAACATGTCTCGCGCGCAGATTTTGCAGCAGGCCGGT
ACTTCCGTTCTGGCGCAGGCTAACCAGGTCCCGCAGAACGTGCTGTCTCTGTTACGTTAA
TAG

SEQ ID NO: 168 FOR PRIMER
AGGCAAGATCTGACACAATATGTATAGGCTACC

SEQ ID NO: 169 REV PRIMER
AGTCAGACGCGTCTATTAACGTAACAGAGACAGCAC

SEQ ID NO: 170 Ha0s.his
GACACAATATGTATAGGCTACCATGCGAACAATTCAACCGACACTGTTGACACAGTACTC
GAGAAGAATGTGACAGTGACACACTCTGTTAACCTGCTCGAAGACAGCCACAACGGAAAA
CTATGTAGATTAAAAGGAATAGCCCCACTACAATTGGGGAAATGTAACATCGCCGGATGG
CTTTTGGGAAACCCAGAATGCGACCCACTGCTTCCAGTGAGATCATGGTCCTACATTGTA
GAAACACCAAACTCTGAGAATGGAATATGTTATCCAGGAGATTTCATCGACTATGAGGAG
CTGAGGGAGCAATTGAGCTCAGTGTCATCATTCGAAAGATTCGAAATATTTCCCAAAGAA
AGCTCATGGCCCAACCACAACACAAACGGAGTAACGGCAGCATGCTCCCATGAGGGGAAA

AGCAGTTTTTACAGAAATTTGCTATGGCTGACGGAGAAGGAGGGCTCATACCCAAAGCTG
AAAAATTCTTATGTGAACAAAAAAGGGAAAGAAGTCCTTGTACTGTGGGGTATTCATCAC
CCGCCTAACAGTAAGGAACAACAGAATCTCTATCAGAATGAAAATGCTTATGTCTCTGTA
GTGACTTCAAATTATAACAGGAGATTTACCCCGGAAATAGCAGAAAGACCCAAAGTAAGA
GATCAAGCTGGGAGGATGAACTATTACTGGACCTTGCTAAAACCCGGAGACACAATAATA
TTTGAGGCAAATGGAAATCTAATAGCACCAATGTATGCTTTCGCACTGAGTAGAGGCTTT
GGGTCCGGCATCATCACCTCAAACGCATCAATGCATGAGTGTAACACGAAGTGTCAAACA
CCCCTGGGAGCTATAAACAGCAGTCTCCCTTACCAGAATATACACCCAGTCACAATAGGA
GAGTGCCCAAAATACGTCAGGAGTGCCAAATTGAGGATGGTTACAGGACTAAGGAACATT
CCGTCCATTCAATCCAGAGGTCTATTTGGAGCCATTGCCGGTTTTATTGAAGGGGGATGG
ACTGGAATGATAGATGGATGGTATGGTTATCATCATCAGAATGAACAGGGATCAGGCTAT
GCAGCGGATCAAAAAAGCACACAAAATGCCATTAACGGGATTACAAACAAGGTGAACACT
GTTATCGAGAAAATGAACATTCAATTCACAGCTGTGGGTAAAGAATTCAACAAATTAGAA
AAAAGGATGGAAAATTTAAATAAAAAAGTTGATGATGGATTTCTGGACATTTGGACATAT
AATGCAGAATTGTTAGTTCTACTGGAAAATGAAAGGACTCTGGATTTCCATGACTCAAAT
GTGAAGAATCTGTATGAGAAAGTAAAAAGCCAATTAAAGAATAATGCCAAAGAAATCGGA
AATGGATGTTTTGAGTTCTACCACAAGTGTGACAATGAATGCATGGAAAGTGTAAGAAAT
GGGACTTATGATTATCCCAAATATTCAGAAGAGTCAAAGTTGAACAGGGAAAAGGTAGAT
GGAGTGAAATTGGAATCAATGGGGATCTATCAGACGCGTACCGGTCATCATCACCATCAC
CATTGA

SEQ ID NO: 171 FOR PRIMER
ATTCTCCGGCGGCCGCTCGACACAATATGTATAGGCTACC

SEQ ID NO: 172 REV PRIMER
AGTCTTGCGGCCGCCTATTAATGGTGATGGTGATGATGCTGATAGATCCCCATTGATTCC

SEQ ID NO: 173 HA0s
GACACAATATGTATAGGCTACCATGCGAACAATTCAACCGACACTGTTGACACAGTACTC
GAGAAGAATGTGACAGTGACACACTCTGTTAACCTGCTCGAAGACAGCCACAACGGAAAA
CTATGTAGATTAAAAGGAATAGCCCCACTACAATTGGGGAAATGTAACATCGCCGGATGG
CTTTTGGGAAACCCAGAATGCGACCCACTGCTTCCAGTGAGATCATGGTCCTACATTGTA
GAAACACCAAACTCTGAGAATGGAATATGTTATCCAGGAGATTTCATCGACTATGAGGAG
CTGAGGGAGCAATTGAGCTCAGTGTCATCATTCGAAAGATTCGAAATATTTCCCAAAGAA
AGCTCATGGCCCAACCACAACACAAACGGAGTAACGGCAGCATGCTCCCATGAGGGGAAA
AGCAGTTTTTACAGAAATTTGCTATGGCTGACGGAGAAGGAGGGCTCATACCCAAAGCTG
AAAAATTCTTATGTGAACAAAAAAGGGAAAGAAGTCCTTGTACTGTGGGGTATTCATCAC
CCGCCTAACAGTAAGGAACAACAGAATCTCTATCAGAATGAAAATGCTTATGTCTCTGTA
GTGACTTCAAATTATAACAGGAGATTTACCCCGGAAATAGCAGAAAGACCCAAAGTAAGA
GATCAAGCTGGGAGGATGAACTATTACTGGACCTTGCTAAAACCCGGAGACACAATAATA
TTTGAGGCAAATGGAAATCTAATAGCACCAATGTATGCTTTCGCACTGAGTAGAGGCTTT
GGGTCCGGCATCATCACCTCAAACGCATCAATGCATGAGTGTAACACGAAGTGTCAAACA
CCCCTGGGAGCTATAAACAGCAGTCTCCCTTACCAGAATATACACCCAGTCACAATAGGA
GAGTGCCCAAAATACGTCAGGAGTGCCAAATTGAGGATGGTTACAGGACTAAGGAACATT
CCGTCCATTCAATCCAGAGGTCTATTTGGAGCCATTGCCGGTTTTATTGAAGGGGGATGG
ACTGGAATGATAGATGGATGGTATGGTTATCATCATCAGAATGAACAGGGATCAGGCTAT
GCAGCGGATCAAAAAAGCACACAAAATGCCATTAACGGGATTACAAACAAGGTGAACACT
GTTATCGAGAAAATGAACATTCAATTCACAGCTGTGGGTAAAGAATTCAACAAATTAGAA
AAAAGGATGGAAAATTTAAATAAAAAAGTTGATGATGGATTTCTGGACATTTGGACATAT
AATGCAGAATTGTTAGTTCTACTGGAAAATGAAAGGACTCTGGATTTCCATGACTCAAAT

GTGAAGAATCTGTATGAGAAAGTAAAAAGCCAATTAAAGAATAATGCCAAAGAAATCGGA
AATGGATGTTTTGAGTTCTACCACAAGTGTGACAATGAATGCATGGAAAGTGTAAGAAAT
GGGACTTATGATTATCCCAAATATTCAGAAGAGTCAAAGTTGAACAGGGAAAAGGTAGAT
GGAGTGAAATTGGAATCAATGGGGATCTATCAG

SEQ ID NO: 174 FOR PRIMER
AGGCAAGATCTGACACAATATGTATAGGCTACC

SEQ ID NO: 175 REV PRIMER
AGTCAGACGCGTCTATTAACGTAACAGAGACAGCAC

SEQ ID NO: 176 HA0sHis(PR8)
DTICIGYHANNSTDTVDTVLEKNVTVTHSVNLLEDSHNGKLCRLKGIAPLQLGKCNIAGWL
LGNPECDPLLPVRSWSYIVETPNSENGICYPGDFIDYEELREQLSSVSSFERFEIFPKESS
WPNHNTNGVTAACSHEGKSSFYRNLLWLTEKEGSYPKLKNSYVNKKGKEVLVLWGIHHPPN
SKEQQNLYQNENAYVSVVTSNYNRRFTPEIAERPKVRDQAGRMNYYWTLLKPGDTIIFEAN
GNLIAPMYAFALSRGFGSGIITSNASMHECNTKCQTPLGAINSSLPYQNIHPVTIGECPKY
VRSAKLRMVTGLRNIPSIQSRGLFGAIAGFIEGGWTGMIDGWYGYHHQNEQGSGYAADQKS
TQNAINGITNKVNTVIEKMNIQFTAVGKEFNKLEKRMENLNKKVDDGFLDIWTYNAELLVL
LENERTLDFHDSNVKNLYEKVKSQLKNNAKEIGNGCFEFYHKCDNECMESVRNGTYDYPKY
SEESKLNREKVDGVKLESMGIYQHHHHHH

SEQ ID NO: 177 STF2Δ.HA0s(PR8)
MAQVINTNSLSLLTQNNLNKSQSALGTAIERLSSGLRINSAKDDAAGQAIANRFTANIKGL
TQASRNANDGISIAQTTEGALNEINNNLQRVRELAVQSANSTNSQSDLDSIQAEITQRLNE
IDRVSGQTQFNGVKVLAQDNTLTIQVGANDGETIDIDLKQINSQTLGLDSLNVHGAPVDPA
SPWTENPLQKIDAALAQVDALRSDLGAVQNRFNSAITNLGNTVNNLSEARSRIEDSDYATE
VSNMSRAQILQQAGTSVLAQANQVPQNVLSLLREFSRYPAQWRPLDTICIGYHANNSTDTV
DTVLEKNVTVTHSVNLLEDSHNGKLCRLKGIAPLQLGKCNIAGWLLGNPECDPLLPVRSWS
YIVETPNSENGICYPGDFIDYEELREQLSSVSSFERFEIFPKESSWPNHNTNGVTAACSHE
GKSSFYRNLLWLTEKEGSYPKLKNSYVNKKGKEVLVLWGIHHPPNSKEQQNLYQNENAYVS
VVTSNYNRRFTPEIAERPKVRDQAGRMNYYWTLLKPGDTIIFEANGNLIAPMYAFALSRGF
GSGIITSNASMHECNTKCQTPLGAINSSLPYQNIHPVTIGECPKYVRSAKLRMVTGLRNIP
SIQSRGLFGAIAGFIEGGWTGMIDGWYGYHHQNEQGSGYAADQKSTQNAINGITNKVNTVI
EKMNIQFTAVGKEFNKLEKRMENLNKKVDDGFLDIWTYNAELLVLLENERTLDFHDSNVKN
LYEKVKSQLKNNAKEIGNGCFEFYHKCDNECMESVRNGTYDYPKYSEESKLNREKVDGVKL
ESMGIYQ

SEQ ID NO: 178 STF2
ATGGCACAAGTAATCAACACTAACAGTCTGTCGCTGCTGACCCAGAATAACCTGAACAAA
TCCCAGTCCGCACTGGGCACCGCTATCGAGCGTCTGTCTTCTGGTCTGCGTATCAACAGC
GCGAAAGACGATGCGGCAGGTCAGGCGATTGCTAACCGTTTCACCGCGAACATCAAAGGT
CTGACTCAGGCTTCCCGTAACGCTAACGACGGTATCTCCATTGCGCAGACCACTGAAGGC
GCGCTGAACGAAATCAACAACAACCTGCAGCGTGTGCGTGAACTGGCGGTTCAGTCTGCT
AACAGCACCAACTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATCACCCAGCGCCTG
AACGAAATCGACCGTGTATCCGGCCAGACTCAGTTCAACGGCGTGAAAGTCCTGGCGCAG
GACAACACCCTGACCATCCAGGTTGGCGCCAACGACGGTGAAACTATCGATATCGATCTG
AAGCAGATCAACTCTCAGACCCTGGGTCTGGACTCACTGAACGTGCAGAAAGCGTATGAT
GTGAAAGATACAGCAGTAACAACGAAAGCTTATGCCAATAATGGTACTACACTGGATGTA
TCGGGTCTTGATGATGCAGCTATTAAAGCGGCTACGGGTGGTACGAATGGTACGGCTTCT
GTAACCGGTGGTGCGGTTAAATTTGACGCAGATAATAACAAGTACTTTGTTACTATTGGT
GGCTTTACTGGTGCTGATGCCGCCAAAAATGGCGATTATGAAGTTAACGTTGCTACTGAC
GGTACAGTAACCCTTGCGGCTGGCGCAACTAAAACCACAATGCCTGCTGGTGCGACAACT

AAAACAGAAGTACAGGAGTTAAAAGATACACCGGCAGTTGTTTCAGCAGATGCTAAAAAT
GCCTTAATTGCTGGCGGCGTTGACGCTACCGATGCTAATGGCGCTGAGTTGGTCAAAATG
TCTTATACCGATAAAAATGGTAAGACAATTGAAGGCGGTTATGCGCTTAAAGCTGGCGAT
AAGTATTACGCCGCAGATTACGATGAAGCGACAGGAGCAATTAAAGCTAAAACTACAAGT
TATACTGCTGCTGACGGCACTACCAAAACAGCGGCTAACCAACTGGGTGGCGTAGACGGT
AAAACCGAAGTCGTTACTATCGACGGTAAAACCTACAATGCCAGCAAAGCCGCTGGTCAT
GATTTCAAAGCACAACCAGAGCTGGCGGAAGCAGCCGCTAAAACCACCGAAAACCCGCTG
CAGAAAATTGATGCCGCGCTGGCGCAGGTGGATGCGCTGCGCTCTGATCTGGGTGCGGTA
CAAAACCGTTTCAACTCTGCTATCACCAACCTGGGCAATACCGTAAACAATCTGTCTGAA
GCGCGTAGCCGTATCGAAGATTCCGACTACGCGACCGAAGTTTCCAACATGTCTCGCGCG
CAGATTCTGCAGCAGGCCGGTACTTCCGTTCTGGCGCAGGCTAACCAGGTCCCGCAGAAC
GTGCTGTCTCTGTTACGT

SEQ ID No: 179 HA1-1his(PR8)

MKFLVNVALVFMVVYISYIYASHNGKLCRLKGIAPLQLGKCNIAGWLLGNPECDPLLPVR
SWSYIVETPNSENGICYPGDFIDYEELREQLSSVSSFERFEIFPKESSWPNHNTNGVTAA
CSHEGKSSFYRNLLWLTEKEGSYPKLKNSYVNKKGKEVLVLWGIHHPPNSKEQQNLYQNE
NAYVSVVTSNYNRRFTPEIAERPKVRDQAGRMNYYWTLLKPGDTIIFEANGNLIAPMYAF
ALSRGFGSGIITSNASMHECNTKCQTPLGAINSSLPYQNIHPVTIGECPKYVRHHHHHH

SEQ ID No: 180 HA1-1his(VN)

MKFLVNVALVFMVVYISYIYAEKKHNGKLCDLDGVKPLILRDCSVAGWLLGNPMCDEFIN
VPEWSYIVEKANPVNDLCYPGDFNDYEELKHLLSRINHFEKIQIIPKSSWSSHEASLGVS
SACPYQGKSSFFRNVVWLIKKNSTYPTIKRSYNNTNQEDLLVLWGIHHPNDAAEQTKLYQ
NPTTYISVGTSTLNQRLVPRIATRSKVNGQSGRMEFFWTILKPNDAINFESNGNFIAPEY
AYKIVKKGDSTIMKSELEYGNCNTKCQTPMGAINSSMPFHNIHPLTIGECPKYVKHHHHH
H

SEQ ID No: 181 HA1-1his(IND)

MKFLVNVALVFMVVYISYIYAEKTHNGKLCDLDGVKPLILRDCSVAGWLLGNPMCDEFIN
VPEWSYIVEKANPTNDLCYPGSFNDYEELKHLLSRINHFEKIQIIPKSSWSDHEASSGVS
SACPYLGSPSFFRNVVWLIKKNSTYPTIKKSYNNTNQEDLLVLWGIHHPNDAAEQTRLYQ
NPTTYISIGTSTLNQRLVPKIATRSKVNGQSGRMEFFWTILKPNDAINFESNGNFIAPEY
AYKIVKKGDSAIMKSELEYGNCNTKCQTPMGAINSSMPFHNIHPLTIGECPKYVKHHHHH
H

SEQ ID No: 182 HA1-1his(NC)

MKFLVNVALVFMVVYISYIYASHNGKLCLLKGIAPLQLGNCSVAGWILGNPECELLISKE
SWSYIVETPNPENGTCYPGYFADYEELREQLSSVSSFERFEIFPKESSWPNHTVTGVSAS
CSHNGKSSFYRNLLWLTGKNGLYPNLSKSYVNNKEKEVLVLWGVHHPPNIGNQRALYHTE
NAYVSVVSSHYSRRFTPEIAKRPKVRDQEGRINYYWTLLEPGDTIIFEANGNLIAPWYAF
ALSRGFGSGIITSNAPMDECDAKCQTPQGAINSSLPFQNVHPVTIGECPKYVRHHHHHH

SEQ ID No: 183 HA1-1(MAL)

ACCACTACCCCAACCAAATCTCACTTCGCAAACCTGAAAGGCACTGAAACCCGTGGCAAG
CTGTGTCCGAAGTGTCTGAACTGCACCGATCTGGACGTCGCACTGGGTCGTCCGAAATGT
ACTGGTAACATTCCGTCCGCGCGTGTCTCCATCCTGCATGAAGTGCGTCCAGTGACCTCC
GGCTGTTTTCCGATTATGCATGATCGTACTAAAATCCGTCAGCTGCCGAACCTGCTGCGT
GGTTACGAACACATTCGTCTGTCCACCCATAACGTTATCAACGCGGAAAACGCGCCGGGC
GGTAGCTATAAAATCGGTACCTCTGGTTCTTGCCCGAACGTGACTAACGGTAACGGCTTC
TTTGCAACCATGGCCTGGGCGGTCCCGAAAAACGACAACAACAAGACCGCGACCAATTCC
CTGACCATCGAAGTCCCGTATATCTGCACCGAAGGTGAAGATCAAATCACGGTTTGGGGC

TTCCACTCCGACAACGAGGCACAAATGGCGAAACTGTACGGTGACAGCAAACCGCAAAAA
TTCACTAGCTCCGCTAACGGTGTTACCACCCACTACGTTTCCCAGATCGGTGGTTTCCCA
AACCAGACCGAAGATGGTGGTCTGCCGCAGTCCGGTCGCATCGTTGTAGATTATATGGTG
CAGAAAAGCGGTAAAACCGGTACCATCACCTACCAGCGTGGCATCCTGCTGCCGCAGAAA
GTTTGGTGCGCTTCCGGTCGTAGCAAAGTAATCAAAGGTTCCCTGCCGCTGATCGGTGAA
GCAGACTGCCTGCACGAGAAATACGGCGGTCTGAACAAAAGCAAGCCGTACTATACCGGC
GAACATGCGAAAGCAATTGGTAACTGTCCAATTTGGGTGAAATAGTAG

SEQ ID No: 184 STF2.HA1-1(MAL)

ATGGCACAAGTAATCAACACTAACAGTCTGTCGCTGCTGACCCAGAATAACCTGAACAAA
TCCCAGTCCGCACTGGGCACCGCTATCGAGCGTCTGTCTTCTGGTCTGCGTATCAACAGC
GCGAAAGACGATGCGGCAGGTCAGGCGATTGCTAACCGTTTCACCGCGAACATCAAAGGT
CTGACTCAGGCTTCCCGTAACGCTAACGACGGTATCTCCATTGCGCAGACCACTGAAGGC
GCGCTGAACGAAATCAACAACAACCTGCAGCGTGTGCGTGAACTGGCGGTTCAGTCTGCT
AACAGCACCAACTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATCACCCAGCGCCTG
AACGAAATCGACCGTGTATCCGGCCAGACTCAGTTCAACGGCGTGAAAGTCCTGGCGCAG
GACAACACCCTGACCATCCAGGTTGGCGCCAACGACGGTGAAACTATCGATATCGATCTG
AAGCAGATCAACTCTCAGACCCTGGGTCTGGACTCACTGAACGTGCAGAAAGCGTATGAT
GTGAAAGATACAGCAGTAACAACGAAAGCTTATGCCAATAATGGTACTACACTGGATGTA
TCGGGTCTTGATGATGCAGCTATTAAAGCGGCTACGGGTGGTACGAATGGTACGGCTTCT
GTAACCGGTGGTGCGGTTAAATTTGACGCAGATAATAACAAGTACTTTGTTACTATTGGT
GGCTTTACTGGTGCTGATGCCGCCAAAAATGGCGATTATGAAGTTAACGTTGCTACTGAC
GGTACAGTAACCCTTGCGGCTGGCGCAACTAAAACCACAATGCCTGCTGGTGCGACAACT
AAAACAGAAGTACAGGAGTTAAAAGATACACCGGCAGTTGTTTCAGCAGATGCTAAAAAT
GCCTTAATTGCTGGCGGCGTTGACGCTACCGATGCTAATGGCGCTGAGTTGGTCAAAATG
TCTTATACCGATAAAAATGGTAAGACAATTGAAGGCGGTTATGCGCTTAAAGCTGGCGAT
AAGTATTACGCCGCAGATTACGATGAAGCGACAGGAGCAATTAAAGCTAAAACTACAAGT
TATACTGCTGCTGACGGCACTACCAAAACAGCGGCTAACCAACTGGGTGGCGTAGACGGT
AAAACCGAAGTCGTTACTATCGACGGTAAAACCTACAATGCCAGCAAAGCCGCTGGTCAT
GATTTCAAAGCACAACCAGAGCTGGCGGAAGCAGCCGCTAAAACCACCGAAAACCCGCTG
CAGAAAATTGATGCCGCGCTGGCGCAGGTGGATGCGCTGCGCTCTGATCTGGGTGCGGTA
CAAAACCGTTTCAACTCTGCTATCACCAACCTGGGCAATACCGTAAACAATCTGTCTGAA
GCGCGTAGCCGTATCGAAGATTCCGACTACGCGACCGAAGTTTCCAACATGTCTCGCGCG
CAGATTCTGCAGCAGGCCGGTACTTCCGTTCTGGCGCAGGCTAACCAGGTCCCGCAGAAC
GTGCTGTCTCTGTTAGCGACCACTACCCCAACCAAATCTCACTTCGCAAACCTGAAAGGC
ACTGAAACCCGTGGCAAGCTGTGTCCGAAGTGTCTGAACTGCACCGATCTGGACGTCGCA
CTGGGTCGTCCGAAATGTACTGGTAACATTCCGTCCGCGCGTGTCTCCATCCTGCATGAA
GTGCGTCCAGTGACCTCCGGCTGTTTTCCGATTATGCATGATCGTACTAAAATCCGTCAG
CTGCCGAACCTGCTGCGTGGTTACGAACACATTCGTCTGTCCACCCATAACGTTATCAAC
GCGGAAAACGCGCCGGGCGGTAGCTATAAAATCGGTACCTCTGGTTCTTGCCCGAACGTG
ACTAACGGTAACGGCTTCTTTGCAACCATGGCCTGGGCGGTCCCGAAAAACGACAACAAC
AAGACCGCGACCAATTCCCTGACCATCGAAGTCCCGTATATCTGCACCGAAGGTGAAGAT
CAAATCACGGTTTGGGGCTTCCACTCCGACAACGAGGCACAAATGGCGAAACTGTACGGT
GACAGCAAACCGCAAAAATTCACTAGCTCCGCTAACGGTGTTACCACCCACTACGTTTCC
CAGATCGGTGGTTTCCCAAACCAGACCGAAGATGGTGGTCTGCCGCAGTCCGGTCGCATC
GTTGTAGATTATATGGTGCAGAAAAGCGGTAAAACCGGTACCATCACCTACCAGCGTGGC
ATCCTGCTGCCGCAGAAAGTTTGGTGCGCTTCCGGTCGTAGCAAAGTAATCAAAGGTTCC
CTGCCGCTGATCGGTGAAGCAGACTGCCTGCACGAGAAATACGGCGGTCTGAACAAAAGC
AAGCCGTACTATACCGGCGAACATGCGAAAGCAATTGGTAACTGTCCAATTTGGGTGAAA
TAGTAG

SEQ ID No: 185 HA1-2(MAL)

AAAGGCACTGAAACCCGTGGCAAGCTGTGTCCGAAGTGTCTGAACTGCACCGATCTGGAC
GTCGCACTGGGTCGTCCGAAATGTACTGGTAACATTCCGTCCGCGCGTGTCTCCATCCTG
CATGAAGTGCGTCCAGTGACCTCCGGCTGTTTTCCGATTATGCATGATCGTACTAAAATC
CGTCAGCTGCCGAACCTGCTGCGTGGTTACGAACACATTCGTCTGTCCACCCATAACGTT
ATCAACGCGGAAAACGCGCCGGGCGGTAGCTATAAAATCGGTACCTCTGGTTCTTGCCCG
AACGTGACTAACGGTAACGGCTTCTTTGCAACCATGGCCTGGGCGGTCCCGAAAAACGAC
AACAACAAGACCGCGACCAATTCCCTGACCATCGAAGTCCCGTATATCTGCACCGAAGGT
GAAGATCAAATCACGGTTTGGGGCTTCCACTCCGACAACGAGGCACAAATGGCGAAACTG
TACGGTGACAGCAAACCGCAAAAATTCACTAGCTCCGCTAACGGTGTTACCACCCACTAC
GTTTCCCAGATCGGTGGTTTCCCAAACCAGACCGAAGATGGTGGTCTGCCGCAGTCCGGT
CGCATCGTTGTAGATTATATGGTGCAGAAAAGCGGTAAAACCGGTACCATCACCTACCAG
CGTGGCATCCTGCTGCCGCAGAAAGTTTGGTGCGCTTCCGGTCGTAGCAAAGTAATCAAA
GGTTGATAG

SEQ ID No: 186 STF2.HA1-2(MAL)

ATGGCACAAGTAATCAACACTAACAGTCTGTCGCTGCTGACCCAGAATAACCTGAACAAA
TCCCAGTCCGCACTGGGCACCGCTATCGAGCGTCTGTCTTCTGGTCTGCGTATCAACAGC
GCGAAAGACGATGCGGCAGGTCAGGCGATTGCTAACCGTTTCACCGCGAACATCAAAGGT
CTGACTCAGGCTTCCCGTAACGCTAACGACGGTATCTCCATTGCGCAGACCACTGAAGGC
GCGCTGAACGAAATCAACAACAACCTGCAGCGTGTGCGTGAACTGGCGGTTCAGTCTGCT
AACAGCACCAACTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATCACCCAGCGCCTG
AACGAAATCGACCGTGTATCCGGCCAGACTCAGTTCAACGGCGTGAAAGTCCTGGCGCAG
GACAACACCCTGACCATCCAGGTTGGCGCCAACGACGGTGAAACTATCGATATCGATCTG
AAGCAGATCAACTCTCAGACCCTGGGTCTGGACTCACTGAACGTGCAGAAAGCGTATGAT
GTGAAAGATACAGCAGTAACAACGAAAGCTTATGCCAATAATGGTACTACACTGGATGTA
TCGGGTCTTGATGATGCAGCTATTAAAGCGGCTACGGGTGGTACGAATGGTACGGCTTCT
GTAACCGGTGGTGCGGTTAAATTTGACGCAGATAATAACAAGTACTTTGTTACTATTGGT
GGCTTTACTGGTGCTGATGCCGCCAAAAATGGCGATTATGAAGTTAACGTTGCTACTGAC
GGTACAGTAACCCTTGCGGCTGGCGCAACTAAAACCACAATGCCTGCTGGTGCGACAACT
AAAACAGAAGTACAGGAGTTAAAAGATACACCGGCAGTTGTTTCAGCAGATGCTAAAAAT
GCCTTAATTGCTGGCGGCGTTGACGCTACCGATGCTAATGGCGCTGAGTTGGTCAAAATG
TCTTATACCGATAAAAATGGTAAGACAATTGAAGGCGGTTATGCGCTTAAAGCTGGCGAT
AAGTATTACGCCGCAGATTACGATGAAGCGACAGGAGCAATTAAAGCTAAAACTACAAGT
TATACTGCTGCTGACGGCACTACCAAAACAGCGGCTAACCAACTGGGTGGCGTAGACGGT
AAAACCGAAGTCGTTACTATCGACGGTAAAACCTACAATGCCAGCAAAGCCGCTGGTCAT
GATTTCAAAGCACAACCAGAGCTGGCGGAAGCAGCCGCTAAAACCACCGAAAACCCGCTG
CAGAAAATTGATGCCGCGCTGGCGCAGGTGGATGCGCTGCGCTCTGATCTGGGTGCGGTA
CAAAACCGTTTCAACTCTGCTATCACCAACCTGGGCAATACCGTAAACAATCTGTCTGAA
GCGCGTAGCCGTATCGAAGATTCCGACTACGCGACCGAAGTTTCCAACATGTCTCGCGCG
CAGATTCTGCAGCAGGCCGGTACTTCCGTTCTGGCGCAGGCTAACCAGGTCCCGCAGAAC
GTGCTGTCTCTGTTAGCGAAAGGCACTGAAACCCGTGGCAAGCTGTGTCCGAAGTGTCTG
AACTGCACCGATCTGGACGTCGCACTGGGTCGTCCGAAATGTACTGGTAACATTCCGTCC
GCGCGTGTCTCCATCCTGCATGAAGTGCGTCCAGTGACCTCCGGCTGTTTTCCGATTATG
CATGATCGTACTAAAATCCGTCAGCTGCCGAACCTGCTGCGTGGTTACGAACACATTCGT
CTGTCCACCCATAACGTTATCAACGCGGAAAACGCGCCGGGCGGTAGCTATAAAATCGGT
ACCTCTGGTTCTTGCCCGAACGTGACTAACGGTAACGGCTTCTTTGCAACCATGGCCTGG
GCGGTCCCGAAAAACGACAACAACAAGACCGCGACCAATTCCCTGACCATCGAAGTCCCG
TATATCTGCACCGAAGGTGAAGATCAAATCACGGTTTGGGGCTTCCACTCCGACAACGAG
GCACAAATGGCGAAACTGTACGGTGACAGCAAACCGCAAAAATTCACTAGCTCCGCTAAC

GGTGTTACCACCCACTACGTTTCCCAGATCGGTGGTTTCCCAAACCAGACCGAAGATGGT
GGTCTGCCGCAGTCCGGTCGCATCGTTGTAGATTATATGGTGCAGAAAAGCGGTAAAACC
GGTACCATCACCTACCAGCGTGGCATCCTGCTGCCGCAGAAAGTTTGGTGCGCTTCCGGT
CGTAGCAAAGTAATCAAAGGTTGATAG

SEQ ID No: 187 HA1-2(SH)

AAAGGCACTCGTACCCGCGGTAAGCTGTGCCCGGACTGCCTGAACTGTACCGATCTGGAT
GTTGCACTGGGTCGTCCGATGTGCGTTGGTACCACCCCGTCTGCGAAAGCCAGCATCCTG
CACGAAGTTCGCCCGGTTACTTCCGGTTGTTTCCCGATTATGCATGATCGTACCAAAATT
CGTCAGCTGCCAAACCTGCTGCGTGGCTATGAAAACATTCGTCTGTCCACTCAAAACGTA
ATCGATGCAGAAAAAGCGCTGGGTGGCCCGTATCGTCTGGGTACCAGCGGCTCCTGCCCG
AACGCGACGAGCAAAAGCGGCTTCTTCGCCACCATGGCTTGGGCCGTTCCGAAAGACAAC
AACAAAAACGCTACGAACCCGCTGACCGTCGAAGTCCCGTACATCTGCACCGAAGGCGAA
GATCAGATCACTGTGTGGGGCTTCCACAGCGATGATAAGACCCAGATGAAAAATCTGTAC
GGTGACTCCAACCCGCAGAAATTCACCTCTTCTGCTAACGGTGTAACGACCCACTACGTT
TCTCAGATCGGTGGTTTCCCGGACCAGACGGAAGATGGCGGTCTGCCTCAGTCCGGCCGC
ATCGTAGTTGATTACATGGTCCAGAAACCGGGTAAGACTGGTACCATTGTTTACCAGCGT
GGTGTACTGCTGCCGCAGAAGGTCTGGTGTGCTTCCGGCCGTTCCAAGGTCATTAAGGGC
TGATA

SEQ ID No: 188 STF2.HA1-2(SH)

ATGGCACAAGTAATCAACACTAACAGTCTGTCGCTGCTGACCCAGAATAACCTGAACAAA
TCCCAGTCCGCACTGGGCACCGCTATCGAGCGTCTGTCTTCTGGTCTGCGTATCAACAGC
GCGAAAGACGATGCGGCAGGTCAGGCGATTGCTAACCGTTTCACCGCGAACATCAAAGGT
CTGACTCAGGCTTCCCGTAACGCTAACGACGGTATCTCCATTGCGCAGACCACTGAAGGC
GCGCTGAACGAAATCAACAACAACCTGCAGCGTGTGCGTGAACTGGCGGTTCAGTCTGCT
AACAGCACCAACTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATCACCCAGCGCCTG
AACGAAATCGACCGTGTATCCGGCCAGACTCAGTTCAACGGCGTGAAAGTCCTGGCGCAG
GACAACACCCTGACCATCCAGGTTGGCGCCAACGACGGTGAAACTATCGATATCGATCTG
AAGCAGATCAACTCTCAGACCCTGGGTCTGGACTCACTGAACGTGCAGAAAGCGTATGAT
GTGAAAGATACAGCAGTAACAACGAAAGCTTATGCCAATAATGGTACTACACTGGATGTA
TCGGGTCTTGATGATGCAGCTATTAAAGCGGCTACGGGTGGTACGAATGGTACGGCTTCT
GTAACCGGTGGTGCGGTTAAATTTGACGCAGATAATAACAAGTACTTTGTTACTATTGGT
GGCTTTACTGGTGCTGATGCCGCCAAAAATGGCGATTATGAAGTTAACGTTGCTACTGAC
GGTACAGTAACCCTTGCGGCTGGCGCAACTAAAACCACAATGCCTGCTGGTGCGACAACT
AAAACAGAAGTACAGGAGTTAAAAGATACACCGGCAGTTGTTTCAGCAGATGCTAAAAAT
GCCTTAATTGCTGGCGGCGTTGACGCTACCGATGCTAATGGCGCTGAGTTGGTCAAAATG
TCTTATACCGATAAAAATGGTAAGACAATTGAAGGCGGTTATGCGCTTAAAGCTGGCGAT
AAGTATTACGCCGCAGATTACGATGAAGCGACAGGAGCAATTAAAGCTAAAACTACAAGT
TATACTGCTGCTGACGGCACTACCAAAACAGCGGCTAACCAACTGGGTGGCGTAGACGGT
AAAACCGAAGTCGTTACTATCGACGGTAAAACCTACAATGCCAGCAAAGCCGCTGGTCAT
GATTTCAAAGCACAACCAGAGCTGGCGGAAGCAGCCGCTAAAACCACCGAAAACCCGCTG
CAGAAAATTGATGCCGCGCTGGCGCAGGTGGATGCGCTGCGCTCTGATCTGGGTGCGGTA
CAAAACCGTTTCAACTCTGCTATCACCAACCTGGGCAATACCGTAAACAATCTGTCTGAA
GCGCGTAGCCGTATCGAAGATTCCGACTACGCGACCGAAGTTTCCAACATGTCTCGCGCG
CAGATTCTGCAGCAGGCCGGTACTTCCGTTCTGGCGCAGGCTAACCAGGTCCCGCAGAAC
GTGCTGTCTCTGTTAGCGAAAGGCACTCGTACCCGCGGTAAGCTGTGCCCGGACTGCCTG
AACTGTACCGATCTGGATGTTGCACTGGGTCGTCCGATGTGCGTTGGTACCACCCCGTCT
GCGAAAGCCAGCATCCTGCACGAAGTTCGCCCGGTTACTTCCGGTTGTTTCCCGATTATG
CATGATCGTACCAAAATTCGTCAGCTGCCAAACCTGCTGCGTGGCTATGAAAACATTCGT

CTGTCCACTCAAAACGTAATCGATGCAGAAAAAGCGCTGGGTGGCCCGTATCGTCTGGGT
ACCAGCGGCTCCTGCCCGAACGCGACGAGCAAAAGCGGCTTCTTCGCCACCATGGCTTGG
GCCGTTCCGAAAGACAACAACAAAAACGCTACGAACCCGCTGACCGTCGAAGTCCCGTAC
ATCTGCACCGAAGGCGAAGATCAGATCACTGTGTGGGGCTTCCACAGCGATGATAAGACC
CAGATGAAAAATCTGTACGGTGACTCCAACCCGCAGAAATTCACCTCTTCTGCTAACGGT
GTAACGACCCACTACGTTTCTCAGATCGGTGGTTTCCCGGACCAGACGGAAGATGGCGGT
CTGCCTCAGTCCGGCCGCATCGTAGTTGATTACATGGTCCAGAAACCGGGTAAGACTGGT
ACCATTGTTTACCAGCGTGGTGTACTGCTGCCGCAGAAGGTCTGGTGTGCTTCCGGCCGT
TCCAAGGTCATTAAGGGCTGATAG

SEQ ID No: 189 HA1-2(Lee)

AAAGGCACTCAGACCCGTGGCAAGCTGTGTCCGAACTGTTTCAACTGCACCGATCTGGAC
GTTGCACTGGGTCGTCCGAAATGCATGGGTAACATCCCGTCTGCGAAGGTAAGCATCCTG
CACGAAGTTAAACCGGTAACCAGCGGCTGTTTCCCGATCATGCACGACAAAACTAAAATT
CGTCAGCTGCCGAACCTGCTGCGTGGTTATGAGAACATTCGTCTGTCTACCTCTAATGTT
ATCAACGCGGAGACTGCACCAGGTGGCCCATACAAAGTAGGTACCAGCGGTTCCTGTCCG
AACGTTGCGAATCGTAACGGCTTCTTCAACACTATGGCGTGGGTTATCCCGAAAGATAAC
AATAAAACTGCAATTAACCCGGTAACTGTAGAAGTTCCGTACATCTGCTCCGAAGGCGAG
GACCAGATTACGGTATGGGGCTTTCACAGCGACGATAAAACCCAGATGGAGCGTCTGTAC
GGTGACTCTAACCCGCAGAAATTCACCTCCTCCGCGAACGGCGTTACCACCCACTATGTT
TCTCAGATCGGCGGTTTCCCGAATCAGACCGAAGACGAAGGCCTGAAGCAGTCCGGCCGT
ATTGTTGTAGACTACATGGTTCAGAAGCCGGGCAAAACTGGTACCATTGTATACCAGCGC
GGCATCCTGCTGCCGCAGAAAGTTTGGTGCGCTTCCGGTCGTAGCAAAGTAATCAAAGGT
TGATAG

SEQ ID No: 190 STF2.HA1-2(Lee)

ATGGCACAAGTAATCAACACTAACAGTCTGTCGCTGCTGACCCAGAATAACCTGAACAAA
TCCCAGTCCGCACTGGGCACCGCTATCGAGCGTCTGTCTTCTGGTCTGCGTATCAACAGC
GCGAAAGACGATGCGGCAGGTCAGGCGATTGCTAACCGTTTCACCGCGAACATCAAAGGT
CTGACTCAGGCTTCCCGTAACGCTAACGACGGTATCTCCATTGCGCAGACCACTGAAGGC
GCGCTGAACGAAATCAACAACAACCTGCAGCGTGTGCGTGAACTGGCGGTTCAGTCTGCT
AACAGCACCAACTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATCACCCAGCGCCTG
AACGAAATCGACCGTGTATCCGGCCAGACTCAGTTCAACGGCGTGAAAGTCCTGGCGCAG
GACAACACCCTGACCATCCAGGTTGGCGCCAACGACGGTGAAACTATCGATATCGATCTG
AAGCAGATCAACTCTCAGACCCTGGGTCTGGACTCACTGAACGTGCAGAAAGCGTATGAT
GTGAAAGATACAGCAGTAACAACGAAAGCTTATGCCAATAATGGTACTACACTGGATGTA
TCGGGTCTTGATGATGCAGCTATTAAAGCGGCTACGGGTGGTACGAATGGTACGGCTTCT
GTAACCGGTGGTGCGGTTAAATTTGACGCAGATAATAACAAGTACTTTGTTACTATTGGT
GGCTTTACTGGTGCTGATGCCGCCAAAAATGGCGATTATGAAGTTAACGTTGCTACTGAC
GGTACAGTAACCCTTGCGGCTGGCGCAACTAAAACCACAATGCCTGCTGGTGCGACAACT
AAAACAGAAGTACAGGAGTTAAAAGATACACCGGCAGTTGTTTCAGCAGATGCTAAAAAT
GCCTTAATTGCTGGCGGCGTTGACGCTACCGATGCTAATGGCGCTGAGTTGGTCAAAATG
TCTTATACCGATAAAAATGGTAAGACAATTGAAGGCGGTTATGCGCTTAAAGCTGGCGAT
AAGTATTACGCCGCAGATTACGATGAAGCGACAGGAGCAATTAAAGCTAAAACTACAAGT
TATACTGCTGCTGACGGCACTACCAAAACAGCGGCTAACCAACTGGGTGGCGTAGACGGT
AAAACCGAAGTCGTTACTATCGACGGTAAAACCTACAATGCCAGCAAAGCCGCTGGTCAT
GATTTCAAAGCACAACCAGAGCTGGCGGAAGCAGCCGCTAAAACCACCGAAAACCCGCTG
CAGAAAATTGATGCCGCGCTGGCGCAGGTGGATGCGCTGCGCTCTGATCTGGGTGCGGTA
CAAAACCGTTTCAACTCTGCTATCACCAACCTGGGCAATACCGTAAACAATCTGTCTGAA
GCGCGTAGCCGTATCGAAGATTCCGACTACGCGACCGAAGTTTCCAACATGTCTCGCGCG
CAGATTCTGCAGCAGGCCGGTACTTCCGTTCTGGCGCAGGCTAACCAGGTCCCGCAGAAC
GTGCTGTCTCTGTTAGCGAAAGGCACTCAGACCCGTGGCAAGCTGTGTCCGAACTGTTTC
AACTGCACCGATCTGGACGTTGCACTGGGTCGTCCGAAATGCATGGGTAACATCCCGTCT

GCGAAGGTAAGCATCCTGCACGAAGTTAAACCGGTAACCAGCGGCTGTTTCCCGATCATG
CACGACAAAACTAAAATTCGTCAGCTGCCGAACCTGCTGCGTGGTTATGAGAACATTCGT
CTGTCTACCTCTAATGTTATCAACGCGGAGACTGCACCAGGTGGCCCATACAAAGTAGGT
ACCAGCGGTTCCTGTCCGAACGTTGCGAATCGTAACGGCTTCTTCAACACTATGGCGTGG
GTTATCCCGAAAGATAACAATAAAACTGCAATTAACCCGGTAACTGTAGAAGTTCCGTAC
ATCTGCTCCGAAGGCGAGGACCAGATTACGGTATGGGGCTTTCACAGCGACGATAAAACC
CAGATGGAGCGTCTGTACGGTGACTCTAACCCGCAGAAATTCACCTCCTCCGCGAACGGC
GTTACCACCCACTATGTTTCTCAGATCGGCGGTTTCCCGAATCAGACCGAAGACGAAGGC
CTGAAGCAGTCCGGCCGTATTGTTGTAGACTACATGGTTCAGAAGCCGGGCAAAACTGGT
ACCATTGTATACCAGCGCGGCATCCTGCTGCCGCAGAAAGTTTGGTGCGCTTCCGGTCGT
AGCAAAGTAATCAAAGGTTGATAG

SEQ ID No: 191 HA1-2(Ohio)

AAAGGCACTAAAACCCGTGGCAAGCTGTGTCCGAAGTGTCTGAACTGCACCGATCTGGAC
GTCGCACTGGGTCGTCCGAAATGTACTGGTAACATTCCGTCCGCGGAAGTCTCCATCCTG
CATGAAGTGCGTCCAGTGACCTCCGGCTGTTTTCCGATTATGCATGATCGTACTAAAATC
CGTCAGCTGCCGAACCTGCTGCGTGGTTACGAACACATTCGTCTGTCCACCCATAACGTT
ATCAACGCGGAAAAGGCGCCGGGCGGTCCCTATAAAATCGGTACCTCTGGTTCTTGCCCG
AACGTGACTAACGGTAACGGCTTCTTTGCAACCATGGCCTGGGCGGTCCCGAAAAACGAC
AACAACAAGACCGCGACCAATTCCCTGACCATCGAAGTCCCGTATATCTGCACCGAAGGT
GAAGATCAAATCACGATTTGGGGCTTCCACTCCGACAGCGAGACACAAATGGCGAAACTG
TACGGTGACAGCAAACCGCAAAAATTCACTAGCTCCGCTAACGGTGTTACCACCCACTAC
GTTTCCCAGATCGGTGGTTTCCCAAACCAGACCGAAGATGGTGGTCTGCCGCAGTCCGGT
CGCATCGTTGTAGATTATATGGTGCAGAAAAGCGGTAAAACCGGTACCATCACCTACCAG
CGTGGCATCCTGCTGCCGCAGAAAGTTTGGTGCGCTTCCGGTCGTAGCAAAGTAATCAAA
GGTTGATAG

SEQ ID No: 192 STF2.HA1-2(Ohio)

ATGGCACAAGTAATCAACACTAACAGTCTGTCGCTGCTGACCCAGAATAACCTGAACAAA
TCCCAGTCCGCACTGGGCACCGCTATCGAGCGTCTGTCTTCTGGTCTGCGTATCAACAGC
GCGAAAGACGATGCGGCAGGTCAGGCGATTGCTAACCGTTTCACCGCGAACATCAAAGGT
CTGACTCAGGCTTCCCGTAACGCTAACGACGGTATCTCCATTGCGCAGACCACTGAAGGC
GCGCTGAACGAAATCAACAACAACCTGCAGCGTGTGCGTGAACTGGCGGTTCAGTCTGCT
AACAGCACCAACTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATCACCCAGCGCCTG
AACGAAATCGACCGTGTATCCGGCCAGACTCAGTTCAACGGCGTGAAAGTCCTGGCGCAG
GACAACACCCTGACCATCCAGGTTGGCGCCAACGACGGTGAAACTATCGATATCGATCTG
AAGCAGATCAACTCTCAGACCCTGGGTCTGGACTCACTGAACGTGCAGAAAGCGTATGAT
GTGAAAGATACAGCAGTAACAACGAAAGCTTATGCCAATAATGGTACTACACTGGATGTA
TCGGGTCTTGATGATGCAGCTATTAAAGCGGCTACGGGTGGTACGAATGGTACGGCTTCT
GTAACCGGTGGTGCGGTTAAATTTGACGCAGATAATAACAAGTACTTTGTTACTATTGGT
GGCTTTACTGGTGCTGATGCCGCCAAAAATGGCGATTATGAAGTTAACGTTGCTACTGAC
GGTACAGTAACCCTTGCGGCTGGCGCAACTAAAACCACAATGCCTGCTGGTGCGACAACT
AAAACAGAAGTACAGGAGTTAAAAGATACACCGGCAGTTGTTTCAGCAGATGCTAAAAAT
GCCTTAATTGCTGGCGGCGTTGACGCTACCGATGCTAATGGCGCTGAGTTGGTCAAAATG
TCTTATACCGATAAAAATGGTAAGACAATTGAAGGCGGTTATGCGCTTAAAGCTGGCGAT
AAGTATTACGCCGCAGATTACGATGAAGCGACAGGAGCAATTAAAGCTAAAACTACAAGT
TATACTGCTGCTGACGGCACTACCAAAACAGCGGCTAACCAACTGGGTGGCGTAGACGGT
AAAACCGAAGTCGTTACTATCGACGGTAAAACCTACAATGCCAGCAAAGCCGCTGGTCAT
GATTTCAAAGCACAACCAGAGCTGGCGGAAGCAGCCGCTAAAACCACCGAAAACCCGCTG
CAGAAAATTGATGCCGCGCTGGCGCAGGTGGATGCGCTGCGCTCTGATCTGGGTGCGGTA
CAAAACCGTTTCAACTCTGCTATCACCAACCTGGGCAATACCGTAAACAATCTGTCTGAA
GCGCGTAGCCGTATCGAAGATTCCGACTACGCGACCGAAGTTTCCAACATGTCTCGCGCG
CAGATTCTGCAGCAGGCCGGTACTTCCGTTCTGGCGCAGGCTAACCAGGTCCCGCAGAAC

GTGCTGTCTCTGTTAGCGAAAGGCACTAAAACCCGTGGCAAGCTGTGTCCGAAGTGTCTG
AACTGCACCGATCTGGACGTCGCACTGGGTCGTCCGAAATGTACTGGTAACATTCCGTCC
GCGGAAGTCTCCATCCTGCATGAAGTGCGTCCAGTGACCTCCGGCTGTTTTCCGATTATG
CATGATCGTACTAAAATCCGTCAGCTGCCGAACCTGCTGCGTGGTTACGAACACATTCGT
CTGTCCACCCATAACGTTATCAACGCGGAAAAGGCGCCGGGCGGTCCCTATAAAATCGGT
ACCTCTGGTTCTTGCCCGAACGTGACTAACGGTAACGGCTTCTTTGCAACCATGGCCTGG
GCGGTCCCGAAAAACGACAACAACAAGACCGCGACCAATTCCCTGACCATCGAAGTCCCG
TATATCTGCACCGAAGGTGAAGATCAAATCACGATTTGGGGCTTCCACTCCGACAGCGAG
ACACAAATGGCGAAACTGTACGGTGACAGCAAACCGCAAAAATTCACTAGCTCCGCTAAC
GGTGTTACCACCCACTACGTTTCCCAGATCGGTGGTTTCCCAAACCAGACCGAAGATGGT
GGTCTGCCGCAGTCCGGTCGCATCGTTGTAGATTATATGGTGCAGAAAAGCGGTAAAACC
GGTACCATCACCTACCAGCGTGGCATCCTGCTGCCGCAGAAAGTTTGGTGCGCTTCCGGT
CGTAGCAAAGTAATCAAAGGTTGATAG

SEQ ID No: 193 FOR PRIMER
CGGATAACAATTCCCCTCTAG

SEQ ID No: 194 REV PRIMER
CGAAGTGAGATTTGGTTGGAGTAGTGGTCGCTAACAGAGACAGCACGTTC

SEQ ID No: 195 FOR PRIMER
CCCGCAGAACGTGCTGTCTCTGTTAGCGACCACTACCCCAACCAAATCTC

SEQ ID No: 196 REV PRIMER
CTATTTCACCCAAATTGGAC

SEQ ID No: 197 REV PRIMER
CACAGCTTGCCACGGGTTTCAGTGCCTTTCGCTAACAGAGACAGCACGTTC

SEQ ID No: 198 FOR PRIMER
CCCGCAGAACGTGCTGTCTCTGTTAGCGAAAGGCACTGAAACCCGTGGC

SEQ ID No: 199 REV PRIMER
GACTAGACGCTCAGCTATCAACCTTTGATTACTTTGCTACGACC

SEQ ID No: 200 REV PRIMER
CACAGCTTACCGCGGGTACGAGTGCCTTTCGCTAACAGAGACAGCACGTTC

SEQ ID No: 201 FOR PRIMER
CGGGATCCAAAGGCACTCGTACCCGCGGTAAG

SEQ ID No: 202 REV PRIMER
GACTAGACGCTCAGCTATCAGCCCTTAATGACCTTGGAACGGCC

SEQ ID No: 203 REV PRIMER
GCATAGTTTTCCTCTGGTCTGTGTTCCTTTCGCTAACAGAGACAGCACGTTC

SEQ ID No: 204 FOR PRIMER
AAAGGAACACAGACCAGAGG

SEQ ID No: 205 REV PRIMER
CTACTACCCTTTTATTACCTTGCTCC

SEQ ID No: 206 REV PRIMER
CACAGCTTGCCACGGGTTTTAGTGCCTTTCGCTAACAGAGACAGCACGTTC

SEQ ID No: 207 FOR PRIMER
AAAGGCACTAAAACCCGTGGC

SEQ ID No: 208 REV PRIMER
GACTAGACGCTCAGCTATCAACCTTTGATTACTTTGCTACGACC

SEQ ID NO: 209 STF2.HA1-1(MAL)
MAQVINTNSLSLLTQNNLNKSQSALGTAIERLSSGLRINSAKDDAAGQAIANRFTANIKG
LTQASRNANDGISIAQTTEGALNEINNNLQRVRELAVQSANSTNSQSDLDSIQAEITQRL
NEIDRVSGQTQFNGVKVLAQDNTLTIQVGANDGETIDIDLKQINSQTLGLDSLNVQKAYD
VKDTAVTTKAYANNGTTLDVSGLDDAAIKAATGGTNGTASVTGGAVKFDADNNKYFVTIG
GFTGADAAKNGDYEVNVATDGTVTLAAGATKTTMPAGATTKTEVQELKDTPAVVSADAKN
ALIAGGVDATDANGAELVKMSYTDKNGKTIEGGYALKAGDKYYAADYDEATGAIKAKTTS
YTAADGTTKTAANQLGGVDGKTEVVTIDGKTYNASKAAGHDFKAQPELAEAAAKTTENPL
QKIDAALAQVDALRSDLGAVQNRFNSAITNLGNTVNNLSEARSRIEDSDYATEVSNMSRA
QILQQAGTSVLAQANQVPQNVLSLLATTTPTKSHFANLKGTETRGKLCPKCLNCTDLDVA
LGRPKCTGNIPSARVSILHEVRPVTSGCFPIMHDRTKIRQLPNLLRGYEHIRLSTHNVIN
AENAPGGSYKIGTSGSCPNVTNGNGFFATMAWAVPKNDNNKTATNSLTIEVPYICTEGED
QITVWGFHSDNEAQMAKLYGDSKPQKFTSSANGVTTHYVSQIGGFPNQTEDGGLPQSGRI
VVDYMVQKSGKTGTITYQRGILLPQKVWCASGRSKVIKGSLPLIGEADCLHEKYGGLNKS
KPYYTGEHAKAIGNCPIWVK

SEQ ID NO: 210 STF2.HA1-2(MAL)
MAQVINTNSLSLLTQNNLNKSQSALGTAIERLSSGLRINSAKDDAAGQAIANRFTANIKG
LTQASRNANDGISIAQTTEGALNEINNNLQRVRELAVQSANSTNSQSDLDSIQAEITQRL
NEIDRVSGQTQFNGVKVLAQDNTLTIQVGANDGETIDIDLKQINSQTLGLDSLNVQKAYD
VKDTAVTTKAYANNGTTLDVSGLDDAAIKAATGGTNGTASVTGGAVKFDADNNKYFVTIG
GFTGADAAKNGDYEVNVATDGTVTLAAGATKTTMPAGATTKTEVQELKDTPAVVSADAKN
ALIAGGVDATDANGAELVKMSYTDKNGKTIEGGYALKAGDKYYAADYDEATGAIKAKTTS
YTAADGTTKTAANQLGGVDGKTEVVTIDGKTYNASKAAGHDFKAQPELAEAAAKTTENPL
QKIDAALAQVDALRSDLGAVQNRFNSAITNLGNTVNNLSEARSRIEDSDYATEVSNMSRA
QILQQAGTSVLAQANQVPQNVLSLLAKGTETRGKLCPKCLNCTDLDVALGRPKCTGNIPS
ARVSILHEVRPVTSGCFPIMHDRTKIRQLPNLLRGYEHIRLSTHNVINAENAPGGSYKIG
TSGSCPNVTNGNGFFATMAWAVPKNDNNKTATNSLTIEVPYICTEGEDQITVWGFHSDNE
AQMAKLYGDSKPQKFTSSANGVTTHYVSQIGGFPNQTEDGGLPQSGRIVVDYMVQKSGKT
GTITYQRGILLPQKVWCASGRSKVIKG

SEQ ID NO: 211 STF2.HA1-2(SH)
MAQVINTNSLSLLTQNNLNKSQSALGTAIERLSSGLRINSAKDDAAGQAIANRFTANIKG
LTQASRNANDGISIAQTTEGALNEINNNLQRVRELAVQSANSTNSQSDLDSIQAEITQRL
NEIDRVSGQTQFNGVKVLAQDNTLTIQVGANDGETIDIDLKQINSQTLGLDSLNVQKAYD
VKDTAVTTKAYANNGTTLDVSGLDDAAIKAATGGTNGTASVTGGAVKFDADNNKYFVTIG
GFTGADAAKNGDYEVNVATDGTVTLAAGATKTTMPAGATTKTEVQELKDTPAVVSADAKN
ALIAGGVDATDANGAELVKMSYTDKNGKTIEGGYALKAGDKYYAADYDEATGAIKAKTTS
YTAADGTTKTAANQLGGVDGKTEVVTIDGKTYNASKAAGHDFKAQPELAEAAAKTTENPL
QKIDAALAQVDALRSDLGAVQNRFNSAITNLGNTVNNLSEARSRIEDSDYATEVSNMSRA
QILQQAGTSVLAQANQVPQNVLSLLAKGTRTRGKLCPDCLNCTDLDVALGRPMCVGTTPS
AKASILHEVRPVTSGCFPIMHDRTKIRQLPNLLRGYENIRLSTQNVIDAEKALGGPYRLG
TSGSCPNATSKSGFFATMAWAVPKDNNKNATNPLTVEVPYICTEGEDQITVWGFHSDDKT
QMKNLYGDSNPQKFTSSANGVTTHYVSQIGGFPDQTEDGGLPQSGRIVVDYMVQKPGKTG
TIVYQRGVLLPQKVWCASGRSKVIKG

SEQ ID NO: 212 STF2

ATGGCACAAGTAATCAACACTAACAGTCTGTCGCTGCTGACCCAGAATAACCTGAACAAA
TCCCAGTCCGCACTGGGCACCGCTATCGAGCGTCTGTCTTCTGGTCTGCGTATCAACAGC
GCGAAAGACGATGCGGCAGGTCAGGCGATTGCTAACCGTTTCACCGCGAACATCAAAGGT
CTGACTCAGGCTTCCCGTAACGCTAACGACGGTATCTCCATTGCGCAGACCACTGAAGGC
GCGCTGAACGAAATCAACAACAACCTGCAGCGTGTGCGTGAACTGGCGGTTCAGTCTGCT
AACAGCACCAACTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATCACCCAGCGCCTG
AACGAAATCGACCGTGTATCCGGCCAGACTCAGTTCAACGGCGTGAAAGTCCTGGCGCAG
GACAACACCCTGACCATCCAGGTTGGCGCCAACGACGGTGAAACTATCGATATCGATCTG
AAGCAGATCAACTCTCAGACCCTGGGTCTGGACTCACTGAACGTGCAGAAAGCGTATGAT
GTGAAAGATACAGCAGTAACAACGAAAGCTTATGCCAATAATGGTACTACACTGGATGTA
TCGGGTCTTGATGATGCAGCTATTAAAGCGGCTACGGGTGGTACGAATGGTACGGCTTCT
GTAACCGGTGGTGCGGTTAAATTTGACGCAGATAATAACAAGTACTTTGTTACTATTGGT
GGCTTTACTGGTGCTGATGCCGCCAAAAATGGCGATTATGAAGTTAACGTTGCTACTGAC
GGTACAGTAACCCTTGCGGCTGGCGCAACTAAAACCACAATGCCTGCTGGTGCGACAACT
AAAACAGAAGTACAGGAGTTAAAAGATACACCGGCAGTTGTTTCAGCAGATGCTAAAAAT
GCCTTAATTGCTGGCGGCGTTGACGCTACCGATGCTAATGGCGCTGAGTTGGTCAAAATG
TCTTATACCGATAAAAATGGTAAGACAATTGAAGGCGGTTATGCGCTTAAAGCTGGCGAT
AAGTATTACGCCGCAGATTACGATGAAGCGACAGGAGCAATTAAAGCTAAAACTACAAGT
TATACTGCTGCTGACGGCACTACCAAAACAGCGGCTAACCAACTGGGTGGCGTAGACGGT
AAAACCGAAGTCGTTACTATCGACGGTAAAACCTACAATGCCAGCAAAGCCGCTGGTCAT
GATTTCAAAGCACAACCAGAGCTGGCGGAAGCAGCCGCTAAAACCACCGAAAACCCGCTG
CAGAAAATTGATGCCGCGCTGGCGCAGGTGGATGCGCTGCGCTCTGATCTGGGTGCGGTA
CAAAACCGTTTCAACTCTGCTATCACCAACCTGGGCAATACCGTAAACAATCTGTCTGAA
GCGCGTAGCCGTATCGAAGATTCCGACTACGCGACCGAAGTTTCCAACATGTCTCGCGCG
CAGATTCTGCAGCAGGCCGGTACTTCCGTTCTGGCGCAGGCTAACCAGGTCCCGCAGAAC
GTGCTGTCTCTGTTACGT

SEQ ID NO: 213 B/Yamagata/16/18 HA

MKAIIVLLMVVTSNADRICTGITSSNSPHVVKTATQGEVNVTGVIPLTTTPTKSHFANLK
GTKTRGKLCPNCLNCTDLDVALGRPMCMGTIPSAKASILHEVRPVTSGCFPIMHDRTKIR
QLPNLLRGYENIRLSTHNVINAERAPGGPYRLGTSGSCPNVTSRNGFFATMAWAVPRDNK
TATNPLTVEVPYICTKGEDQITVWGFHSDDKTQMKNLYGDSNPQKFTSSANGVTTHYVSQ
IGDFPNQTEDGGLPQSGRIVVDYMVQKPGKTGTIVYQRGVLLPQKVWCASGRSKVIKGSL
PLIGEADCLHEKYGGLNKSKPYYTGEHAKAIGNCPIWVKTPLKLANGTKYRPPAKLLKER
GFFGAIAGFLEGGWEGMIAGWHGYTSHGAHGVAVAADLKSTQEAINKITKNLNSLSELEV
KNLQRLSGAMDELHNEILELDEKVDDLRADTISSQIELAVLLSNEGIINSEDEHLLALER
KLKKMLGPSAVDIGNGCFETKHKCNQTCLDRIAAGTFNAGEFSLPTFDSLNITAASLNDD
GLDNHTILLYYSTAASSLAVTLMIAIFIVYMVSRDNVSCSICL

SEQ ID NO: 787 B/Victoria/2/87 HA

MKAIIVLLMVVTSNADRICTGITSSNSPHVVKTATQGEVNVTGVIPLTTTPTKSHFANLK
GTKTRGKLCPKCLNCTDLDVALGRPKCTGTIPSAKASILHEVKPVTSGCFPIMHDRTKIR
QLPNLLRGYEHIRLSTHNVINAETAPGGPYKVGTSGSCPNVTNGNGFFATMAWAVPKNDN
NKTATNPLTVEVPYICTEGEDQITVWGFHSDNEAQMVKLYGDSKPQKFTSSANGVTTHYV
SQIGGFPNQAEDGGLPQSGRIVVDYMVQKSGKTGTITYQRGILLPQKVWCASGRSKVIKG
SLPLIGEADCLHEKYGGLNKSKPYYTGEHAKAIGNCPIWVKTPLKLANGTKYRPPAKLLK
EKGFFGAIAGFLEGGWEGMIAGWHGYTSHGAHGVAVAADLKSTQEAINKITKNLNSLSEL
EVKNLQRLSGAMDELHNKILELDEKVDDLRADTISSQIELAVLLSNEGIINSEDEHLLAL
ERKLKKMLGPSAVEIGNGCFETKHKCNQTCLDRIAAGTFNAGEFSLPTFDSLNITAASLN
DDGLDNHTILLYYSTAASSLAVTLMIAIFIVYMVSRDNVSCSICL

EXAMPLE 13 FLAGELLIN PROTEINS WITH ENGINEERED CYSTEINE RESIDUES FOR CHEMICAL CONJUGATION

The crystal structure for the bacterial flagellin *S. typhimurium* flagellin type 2 (STF2, SEQ ID NO: 312) has been reported (Yonekura, K., *et al., Nature* 424:643-650(2003)). Complete atomic model of the bacterial flagellar filament by electron cryomicroscopy has been reported, and the most detailed structure-function study of the TLR5 activation was based on this structure (Smith, K.D., *et al., Nature Immunology* 4:1247-1253 (2003)). Toll-like Receptor 5 (TLR5)recognizes a conserved site on flagellin required for protofilament formation and bacterial motility (*Nature Immunology* 4(12):1247-1253). Mutational analysis demonstrated that the TLR5 activity of flagellin resides in two regions in the N- and C-terminal domains, stretches of 39 and 31 amino acids, respectively (Smith, *et al.)*. These regions are highlighted in grey in Figure 18. Alanine-scanning mutagenesis of these regions identified a number of mutations which reduced TLR5 activation, but no single mutation completely abrogated activity, suggesting that there is a degree of flexibility or redundancy in the TLR5 binding site. This makes sense from an evolutionary perspective, otherwise the bacteria could easily evolve to evade detection by TLR5. Deletion of the hypervariable hinge region and retention of the conserved N- and C-terminal domains results in a protein (STF2Δ) that retains full TLR5 activity, demonstrating that the N- and C-terminal domains (together) are necessary and sufficient for TLR5 activation.

Lysine residues are convenient substrates for chemical conjugation of diverse chemical structures to a protein carrier such as flagellin. Of the 30 lysine residues in STF2 (SEQ ID NO: 312), only one lies within the experimentally-defined TLR5 activation site, and this particular lysine is not conserved among all bacterial flagellins (Smith, *et al.*, Nature *Immunology* 4:1247-1253 (2003)). Of the remaining 29 lysines, 24 are located in the hypervariable hinge region of flagellin, thus leaving only 5 lysine residues in STF2Δ (SEQ ID NO: 313). The structure in Figures 19 and 20 shows that most lysines are spatially distal to the TLR5 activation domain, thus it would seem that lysines could be randomly conjugated with peptide or carbohydrate antigens with a reasonable probability of not interfering with the TLR5 activity of

the resulting conjugated STF2Δ (SEQ ID NO: 313) protein. Five lysines appear to border the TLR5 activation site quite closely; these residues would be the first to consider mutating if antigen conjugation to lysine affects TLR5 activity.

Possible additional sites for conjugation (such as cysteines or additional lysines) may be engineered into full-length flagellin (STF2 (SEQ ID NO: 312)) or STF2Δ (SEQ ID NO: 313) by placing such residues in the hypervariable region or in the tail of the conserved N- and C- terminal domains distal to the TLR5 binding site. Chemical methods for conjugation to other amino acids also exist. These include carboxy amino acids (glutamic acid, aspartic acid) and the carboxyl terminal amino acid, arginine, histidine, tryptophan, tyrosine, and serine. Any of these strategies may be utilized to conjugate antigenic structures to flagellin without intefering with the TLR5 activation site.

Some types of antigens including polysaccharides are not amenable to recombinant fusion DNA technology nor to synthetic peptide chemistry, and thus cannot be genetically or synthetically linked to a ligand for a Toll-like Receptor (TLR). In addition, it is possible that genetic or synthetic fusion of a peptide to a TLR ligand may inhibit the proper folding of the TLR ligand. Chemical conjugation of the antigen to the TLR ligand, which is folded and purified, may be useful. The chemical conjugation of peptide antigens also permits the placement of antigen at specific sites on the TLR ligand, thus limiting interference with receptor binding of the TLR ligand or maximizing exposure to antigen receptors. The conjugation of many peptide antigens to a TLR ligand may also maximize immunogenicity by increasing the conformational heterogeneity with which antigen is presented to the immune system. To provide a site for chemical conjugation of such antigens to flagellin, a cysteine residue was engineered at different sites in the gene encoding STF2Δ (flagellin with the hypervariable hinge region deleted; SEQ ID NO: 313 and SEQ ID NO: 314).

Materials and Methods

Polymerase chain reaction (PCR): Platinum® PCR SuperMix High Fidelity kit (catalog number 12532-016, Invitrogen Corporation, Carlsbad, CA) was used for

all PCR amplifications using the following protocol based on the manufacturer's instructions.

1. The following components were added in any order to each reaction tube:
   a. 45 µl Platinum® PCR SuperMix High Fidelity (PCR reaction buffer)
   b. Primer solution (10 pMol final concentration of each primer)
   c. Template DNA solution (10 ng plasmid DNA)
2. Reaction volume was made up to 50 µl with water
3. Tubes were capped and loaded in thermal cycler
4. Tubes were incubated at 94°C for 30 to 120 seconds to completely denature the template and activate the enzyme
5. The following PCR amplification was performed for 25-35 cycles:
   a. Denature at 94°C for 15-30 seconds
   b. Anneal at 50-55°C for 15-30 seconds
   c. Extend at 68-72°C for 1 minute per kb of PCR product size
   d. If necessary, cool to 4°C and hold until ready for next process

Construction of STF2Δ gene (SEQ ID NO: 314): Full length flagellin of *Salmonella typhimurium* fljb (STF2, SEQ ID NO: 312) is encoded by a 1.5 kb gene (SEQ ID NO: 315). To generate STF2Δ (SEQ ID NO: 314), the sequence corresponding to the hypervariable hinge region (amino acids 170 to 415 of SEQ ID NO: 312) was deleted and replaced with a short flexible linker (GAPVDPASPW (SEQ ID NO: 336)) designed to facilitate interactions of the NH2 and COOH terminal sequences of STF2 (SEQ ID NO: 315) necessary for TLR5 signaling (Smith, K. D., *et al.*, (2004). Toll-like Receptor 5 recognizes a conserved site on flagellin required for protofilament formation and bacterial motility (Smith, *et al.*, *Nature Immunology* 4:1247-53 (2003)).

To generate the STF2Δ plasmid, a two-step PCR was used (Figure 21). In the first reaction, a plasmid encoding STF2.OVA (a fusion of full-length STF2 with

full-length ovalbumin, SEQ ID NO: 316) was mixed with primers STF28BGF-1 (SEQ ID NO: 317) and STF28MCR-1 (SEQ ID NO: 318) in PCR reaction buffer, and the mixture was amplified in a PCR reaction as described above. In a parallel reaction, STF2.OVA template plasmid was mixed with primers STF28.MCF-2 (SEQ ID NO: 319) and STF28ECR-2 (SEQ ID NO: 320) in PCR reaction buffer, and the mixture was amplified in a PCR reaction as described above.

The PCR amplification reactions generated ~500 bp and 270 bp fragments, respectively. These PCR products were combined with primers STF28BGF-1 (SEQ ID NO: 317) and STF28ECR-2 (SEQ ID NO: 320) in PCR reaction buffer, and the mixture was amplified in a PCR reaction as described above. The amplified DNA product from this reaction (770 bp) was digested with *Bgl*II and *EcoR*I restriction enzymes for 2 hours at 37°C, and ligated into pMTBiP/V5-His B (Invitrogen, Carlsbad, CA) that had previously been digested with *Bgl*II and *EcoRI* and treated with calf intestinal phosphatase (CIP). An aliquot of the ligation mix was used to transform *E. coli* TOP10 cells. The resultant construct pMT/STF2Δ was used to generate the cysteine modified STF2Δ constructs.

Construction of STF2Δ genes engineered to express a single cysteine residue: To introduce cysteine residues at defined positions in the STF2 polypeptide, pMT/STF2Δ was used as a template in PCR with primer pairs as defined below.

To construct the STF2Δ.3'Cys (SEQ ID NO: 325) plasmid, pMT/STF2Δ plasmid was mixed with 3'Forward1 (SEQ ID NO: 321) and 3'Reverse1 (SEQ ID NO: 322) primers in PCR reaction buffer, and the mixture was amplified in a PCR reaction as described above.

To construct the 5'Cys.STF2Δ (SEQ ID NO: 326) plasmid, pMT/STF2Δ plasmid was mixed with 5'Forward2 (SEQ ID NO: 323) and 5'Reverse2 (SEQ ID NO: 324) primers in PCR reaction buffer, and the mixture was amplified in a PCR reaction as described above.

For STF2Δ.3'Cys (SEQ ID NO: 325) and 5'Cys.STF2Δ (SEQ ID NO: 326) constructs, the PCR products generated were digested with *Nde*1 and *Blp*1 restriction enzymes and purified by agarose gel electrophoresis. The purified fragments were ligated by compatible ends to pET24a (Novagen, Madison, WI) plasmid DNA that had been previously digested with *Nde*1 and *Blp*1 restriction

enzymes and CIP-treated. The resulting constructs are designated pET/STF2Δ.3'Cys and pET/5'Cys.STF2Δ, respectively.

To construct the STF2Δ.HingeCys (SEQ ID NO: 331) plasmid, a two step PCR similar to that described in Figure 18 was employed. One fragment was generated by mixing pMT/STF2Δ with paired HingeForward1 (SEQ ID NO: 327) and HingeReverse3 (SEQ ID NO: 328) primers in PCR reaction buffer, and the mixture was amplified in a PCR reaction as described above. Another fragment was generated by mixing pMT/STF2Δ with paired HingeForward2 (SEQ ID NO: 329) and HingeReverse4 (SEQ ID NO: 330) primers in PCR reaction buffer, and the mixture was amplified in a PCR reaction as described above. Aliquots from each of these PCR reactions were combined and mixed with HingeForward1 (SEQ ID NO: 327) and HingeReverse4 (SEQ ID NO: 330) primers in PCR reaction buffer, and the mixture was amplified in a PCR reaction as described above. The final PCR product was purified by agarose gel electrophoresis and digested with *Nde*1 and *Blp*1, and the purified fragment was ligated into the pET24a vector as described above. The resulting construct is designated pET/STF2Δ.HingeCys.

Protein expression: Plasmids pET/STF2Δ.3'Cys, pET/5'Cys.STF2Δ, and pET/STF2Δ.HingeCys were transformed into competent *E. coli* BLR(DE3) cells as follows. To a 1.5-ml snap-cap polypropylene tubes pre-chilled on ice was added 20 µl aliquot of BLR (DE3) cells and 1 µl plasmid DNA (1 µg/µl), and the mixture was incubated for 30 minutes. The tubes were heated for exactly 30 seconds in a 42°C water bath without shaking, and then placed on ice for an additional 2 minutes. To each tube of heat shocked cells were added 250 µl of room temperature SOC medium.

The cells were recovered at 37°C (shaking at 250 rpm) for 60 minutes prior to plating on selective media containing kanamycin. Various ailquots (50-100 µl) of the transformation mix were plated and the plates were incubatedat 37°C for 15 to 18 hours. Colonies were picked and inoculated into 2 ml Luria–Bertani (LB) broth supplemented with 25 µg/ml kanamycin and 12.5 µg/ml tetracycline, and cultured overnight. Fresh LB cultures were inoculated by diluting an aliquot of the overnight cultures 1:100 and cultured at 37°C with shaking. When the OD_{600} of the culture reached 0.6 to 1.0, protein expression was induced by the addition of isopropyl thio-

β-D-galactoside (IPTG) to a final concentration of 1 mM. Several hours after induction, the cells were harvested for analysis of protein expression by SDS-PAGE. Glycerol stocks were prepared from cultures grown in LB supplemented with 0.5% glucose, 25 μg/ml kanamycin and 12.5 μg/ml tetracycline and were frozen at -80°C following the addition of glycerol (7% final).

Protein purification: Proteins STF2Δ.3'Cys (SEQ ID NO: 332) and STF2Δ.HingeCys (SEQ ID NO: 333) were expressed and purified as follows. Glycerol stocks of *E. coli* BLR(DE3) cells harboring the desired plasmids were inoculated into 10 L shake-flasks containing LB and incubated at 37°C with constant shaking. When cultures reached an optical density of A_{600}=0.8, protein expression was induced by the addition of 1 mM IPTG and cultures were incubated at 37°C with constant shaking for 4 hours before harvesting. Cells were collected from 10 L of culture by low-speed centrifugation at 5,000 rpm (SLA3000 rotor) for 10 minutes and suspended in 50 mM Tris-HCl pH 8.0, 1 mM EDTA, 1 mM DTT (100 ml/10 L).

The cells were disrupted by passing the cell suspension twice through a microfluidizer at 18,000 psi. The insoluble material was separated from soluble proteins by centrifugation at 10,000 rpm (SS34 rotor) for 15 minutes. Under the culture conditions described, all STF2Δ proteins fractionated with the insoluble material and formed stable inclusion bodies that were collected as a solid pellet following centrifugation. The inclusion bodies were washed twice with 50 mM Tris-HCl, pH 8.0, 0.1 M NaCl and 0.5% Triton X-100, followed by two washes with the same buffer without detergent. Each wash was performed using a dounce homogenizer and the cleaned inclusion bodies (IBs) were collected by centrifugation at 10,000 rpm (SS34 rotor) for 15 minutes. The purified IBs were washed a final time with 50 mM Tris-HCl, pH 8.0 and stored as a cell pellet at -80°C until needed.

IB material was thawed and solubilized in 8 M urea at pH 4.0. This step selectively solubilized the target protein while leaving a significant amount of debris and contaminating proteins as a solid precipitate which was removed by centrifugation. Because of the single cysteine present in these proteins, all subsequent purification procedures were carried out using buffers containing 1 mM DTT as a reducing agent. Solubilized protein was captured using SP fast flow sepharose (30 ml, XK16) and selectively eluted with 8 M urea in 25 mM Na Acetate

($C_2H_3O_2Na$), pH 4.0, 1 mM DTT, 1 mM EDTA and 0.2 M NaCl. To eliminate protein precipitation following this step, the pH of the SP elution was adjusted from 4.0 to 8.0 by dialysis against 50 mM Tris-HCl, pH 8.0, 1 mM DTT, 1 mM EDTA before protein refolding. The dialyzed material was refolded by direct dilution (10-fold) into 50 mM Tris-HCl pH 8.0, 1 mM DTT and 1 mM EDTA such that the final protein concentration was less than 0.1 mg/ml. The refolded SP pool was loaded directly onto Q high performance sepharose (30 ml, XK16) and bound protein was eluted with a 20 column volume linear gradient from 0 to 0.5 M NaCl in 50 mM Tris-HCl, pH 8.0, 1 mM DTT, 1 mM EDTA. This chromatography step yielded a single peak that eluted at a conductivity of approximately 15 ms/cm. The eluted material was pooled and stored at -80°C.

Protein characterization: Proteins were characterized for purity, identity, endotoxin content, and biological activity using the following assays.

SDS-PAGE: Proteins (typically 5 μg) were diluted in SDS-PAGE sample buffer (1% SDS, 30 mM Tris-HCl, pH 6.8, 4% glycerol, 0.1 mg/ml bromophenol blue) with and without 5 mM β-mercaptoethanol. The samples were boiled for 5 minutes and loaded onto a 4-20% SDS polyacrylamide gel. Following electrophoresis, gels were stained with coomassie blue to visualize protein bands.

Endotoxin assay: Endotoxin levels were measured using the QCL-1000 Quantitative Chromogenic LAL test kit (BioWhittaker #50-648U), following the manufacturer's instructions for the microplate method.

Protein Assay: Protein concentrations were determined by the MicroBCA Protein Assay Reagent Kit in a 96-well format using BSA as a standard (Pierce Biotechnology), following the manufacturer's instructions.

Flagellin ELISA: Protein integrity and concentration were examined by ELISA with antibodies specific for flagellin. 96-well ELISA plates were coated overnight at 4°C with serial dilutions of each target protein, in PBS starting at 5 μg/ml. Plates were blocked with 200 μl/well of Assay Diluent Buffer (ADB; BD Pharmingen) for one hour at room temperature then washed three times in phosphate-buffered saline containing Tween-20 (PBS-T, 12 mM $NaPO_4$, 137 mM NaCl, 2.7 mM KCl, 0.05% Tween 20). Rabbit polyclonal anti-flagellin antibody diluted in ADB (100 μl/well, 1:5000) was added to all wells and the plates were

incubated for 1 hour at room temperature or overnight at 4°C, then washed three times with PBS-T. HRP-labeled goat anti-rabbit IgG antibody (Jackson Immunochemical) diluted in ADB was added (100 µl/well, 1:5000) and the plates were incubated at room temperature for 1 hour. The plates were washed three times with PBS-T. After adding TMB Ultra substrate (Pierce) and monitoring color development, A_{450} was measured on a Tecan Farcyte microplate spectrophotometer.

TLR5 bioactivity assay: HEK293 cells (ATCC, Cat#CRL-1573, Manassas, VA) constitutively express TLR5, and secrete several soluble factors, including IL-8, in response to TLR5 signaling. Cells were seeded in 96 well microplates (50,000 cells/well), and recombinant test proteins were added. The next day, the conditioned medium was harvested, transferred to a clean 96-well microplate, and frozen at -20°C. After thawing, the conditioned medium was assayed for the presence of IL-8 in a sandwich ELISA using an anti-human IL-8 matched antibody pair (Pierce; Rockford, IL, #M801E and #M802B) following the manufacturer's instructions. Optical density was measured using a microplate spectrophotometer

Results and Discussion

Protein yield and purity: The final yield and endotoxin levels of each protein are shown below.

Protein	SEQ ID NO:	Yield (mg)	Endotoxin (EU/µg)
STF2Δ.3'Cys	332	75.0	0.01
STF2Δ.HingeCys	333	117.0	0.004

Flagellin integrity: Engineering a single cysteine into STF2Δ does not diminish its recognition by flagellin-specific antibodies, as shown in Figure 22.

TLR5 agonist activity in the HEK293 IL-8 assay: Engineering a single cysteine into STF2Δ does not diminish the TLR agonist activity of the protein, as shown in Figure 23. Cells exposed to STF2Δ.3'Cys (SEQ ID NO: 332) or STF2Δ.HingeCys (SEQ ID NO: 333) secreted IL-8 at levels comparable to those induced by exposure to full-length flagellin (STF2, SEQ ID NO: 312) or STF2Δ without cysteine residues (SEQ ID NO: 313).

Fusion of full-length flagellin (STF2, SEQ ID NO: 312) or hinge region-deleted flagellin (STF2Δ, SEQ ID NO: 313) to a protein antigen such as West Nile virus envelope protein or influenza A hemagglutinatin significantly increases the immunogenicity of the fused antigen. This approach is useful for protein or peptide antigens which can be encoded in a genetic fusion construct expressing both flagellin and the antigen of interest in a single protein, but it cannot be applied to non-protein or non-peptide antigens, such as polysaccharide antigens. Since many bacteria and tumor cells express specific polysaccharide structures that are antigenic but poorly immunogenic, it is important to devise scalable methods for increasing the immunogenicity and protective efficacy of such structures. Chemical conjugation to a TLR ligand, such as flagellin, is one potential strategy.

There are several methods of chemical linkage, but one of the simplest is to couple an antigen to a free thiol group such as that on a single, unmatched cysteine residue. Since flagellin does not natively contain any cysteines, cysteine residues were engineered into flagellin at one of three positions: the amino terminus of flagellin (5'Cys.STF2Δ, SEQ ID NO: 326 and SEQ ID NO: 334), the carboxy terminus of flagellin (STF2Δ.3'Cys, SEQ ID NO: 325 and SEQ ID NO: 332), and in the deleted hinge region of flagellin (STF2Δ.HingeCys, SEQ ID NO: 331 and SEQ ID NO: 333). All constructs were produced in the expression vector pET24a for *E. coli* expression. When purified from *E. coli* cells, STF2Δ.HingeCys (SEQ ID NO: 333) and STF2Δ.3'Cys (SEQ ID NO: 332) proteins retain all antigenic and TLR5 biological activity properties of the unmodified STF2Δ protein, thus confirming that the introduction of a single cysteine residue does not negatively impact the expression, purification, refolding, or biological activity of STF2Δ.

CONJUGATION OF INFLUENZA HEMAGGLUTININ MATURATIONAL CLEAVAGE SITE PEPTIDE TO STF2Δ.hingeCYS

A peptide representing the maturational cleavage site of influenza hemagglutinin was chemically conjugated a modified flagellin protein, STF2Δ.HingeCys (SEQ ID NO: 333).

Materials and Methods

Production of STF2Δ.HingeCys protein: STF2Δ.HingeCys protein (SEQ ID NO: 333) was expressed and purified as described in the previous Example.

Synthesis of H1C1 peptides The sequence NH_2-NIPSIQSRGLFFAIAGFIE-COOH (SEQ ID NO: 337) represents the maturational cleavage site of influenza A/H1N1 hemagglutinin (Bianchi, E., et al. (2005). Universal influenza B vaccine based on the maturational cleavage site of the hemagglutinin precursor (Bianchi, *et al.*, *J. Virol.* 79:7380-7388 (2005)). Two peptides were designed with an extra cysteine or an extra lysine on the N-terminus to facilitate chemical linkage to carrier proteins, resulting in the following peptide sequences:

CysH1C1 (SEQ ID NO: 338) NH_2-CNIPSIQSRGLFFAIAGFIE-COOH (SEQ ID NO: 338)

LysH1C1 (SEQ ID NO: 339) NH_2-KNIPSIQSRGLFFAIAGFIE-COOH (SEQ ID NO: 339)

The peptides were synthesized by Anaspec, Inc., (San Jose, CA) utilizing standard Fmoc chemistry, after which they were cleaved from the matrix with trifluoroacetic acid (TFA), purified by reversed-phase HPLC and lyophilized.

Conjugation of CysH1C1 peptide to STF2Δ.HingeCys: STF2Δ.HingeCys protein (SEQ ID NO: 333) at a concentration of 4.9 mg/mL was dialyzed overnight into Buffer A [1x phosphate-buffered saline (PBS), 5 mM EDTA, pH 7.2]. $BM(PEO)_2$ (Pierce Biotechnology, Rockford, IL), a homobifunctional maleimide crosslinker, was dissolved in DMSO (dimethyl sulfoxide) and added to a final concentration of 2.3 mM. After incubating for 1 hour at room temperature, the free crosslinker was removed from the protein using a 5 ml Superdex 25 HiTrap desalting column. CysH1C1 peptide (SEQ ID NO: 338) was dissolved in DMSO and added to the maleimide-derivatized STF2Δ.HingeCys protein at a final concentration of 0.32 mM. After incubating 3 hours at room temperature, the reaction was stopped by adding DTT (dithiothreitol) to a final concentration of 40 mM. The protein and protein-peptide conjugate were separated from free peptide by fractionation on a Superdex 200 10/300 size-exclusion chromatography (SEC) column (GE/Amersham; Piscataway, NJ) equilibrated in 1x Tris-buffered saline (TBS), pH 8.0.

Protein characterization: Proteins were characterized for purity, identity, endotoxin content, and biological activity as described in the previous Example.

Results and Discussion

Characterization of the STF2Δ.HingeCys:CysH1C1 peptide conjugate:

Conjugation of the CysH1C1 peptide (SEQ ID NO: 338) to STF2Δ.HingeCys (SEQ ID NO: 333) was assayed by SDS-PAGE with coomassie staining. Peptide conjugation to the protein resulted in a doublet band in which the faster-migrating band corresponded to the un-conjugated protein and the slower migrating band corresponded to the protein-peptide conjugate.

A cloudy precipitate was seen in the conjugation mixture due to the low solubility of the peptide. To ascertain if the protein-peptide conjugate was preciptitating, the mixture was centrifuged (16,000xg for 15 minutes) and the resulting pellet and supernatant fractions analyzed by SDS-PAGE. The majority of the peptide-conjugated protein remains in the supernatant indicating that it is still soluble.

Size-exclusion chromatography (SEC) of the STF2Δ.HingeCys:CysH1C1 conjugate was performed to separate the conjugated protein from remaining crosslinker and unonjugated peptide. A single major protein peak eluted in the included range of the column at the expected volume for monomeric STF2Δ, while there is very little material eluting in the void (Figure 24), indicating that most of the fractionating protein is monomeric.

The monomeric nature of the STF2Δ.HingeCys:CysH1C1 conjugate was confirmed by SDS-PAGE analysis of the S200 fractions (Figure 25). The protein-peptide conjugate co-elutes with the unconjugated protein in the same ratio as in the load sample, indicating that the conjugate species is a monomer.

Bioactivity of the STF2Δ.HingeCys:CysH1C1 conjugate mixture was found to be equivalent to the unconjugated STF2Δ.HingeCys protein, indicating that the attachment of the $BM(PEO)_2$ crosslinker and the peptide conjugation reaction does not inhibit TLR5 stimulatory activity (Figure 8).

CONJUGATION OF A PAM_3CYS-CONTAINING LIPOPEPTIDE TO AN INFLUENZA HEMAGGLUTININ ANTIGEN

Several types of Toll-like Receptors (TLRs) are potently activated by lipids or lipid-conjugates. These include TLR2/1 (diacyl lipoproteins and GPI-linked proteins), TLR2/6 (triacyl lipoproteins and GPI-linked proteins) and TLR4 (lipopolysaccharides). The utility of vaccines composed of antigens liked to such lipids is obvious: the lipids are strong activators of innate immune pathways and yet are themselves poorly immunogenic. However, the production of such conjugates is very complex. Biosynthesis of recombinant lipoproteins in bacteria is limited by low productivity and extreme difficulty in purifying the resulting protein-lipopeptide fusion. Chemical synthesis of lipopeptide antigens is more straight forward, but is limited to peptide antigens. Here we demonstrate the chemical conjugation of a Pam_3Cys lipopeptide, a TLR2/6 agonist, to influenza HA1-$1His_6$ in a process which bypasses the difficulties inherent in both biosynthesis and complete chemical synthesis of lipidated antigens.

Materials and Methods

Production of HA1-$1His_6$(PR8)Bv (SEQ ID NO: 179) HA1-$1His_6$ (PR8)Bv was prepared as described herein.

Synthesis of $Pam_3CS(K)_4GC$ (SEQ ID NO: 335) Custom lipopeptide synthesis was performed by Anaspec, Inc. (San Jose, CA) using standard solid-phase Fmoc chemistry, after which the lipopeptide was cleaved from the matrix with trifluoroacetic acid (TFA), purified by reversed-phase HPLC and lyophilized.

Conjugation of $Pam_3CS(K)_4GC$ (SEQ ID NO:335) to HA1-$1His_6$(PR8)Bv (SEQ ID NO: 179) HA1-$1His_6$(PR8)Bv (SEQ ID NO: 179) was dialyzed into Buffer A (1x PBS + 5mM EDTA, pH 7.2). Sulfo-SMCC (Pierce; Rockland, IL) a heterobifunctional maleimide/NHS-ester crosslinker was dissolved in DMSO and added to a final concentration of 0.432 mM. After incubating for 30 min. at room temperature the protein was desalted into Buffer A using a G-25 HiTrap desalting column (GE/Amersham; Piscataway, NJ). Triton X-114 (TX-114) (Sigma; St Louis, MO) was added to a final concentration of 1%(w/v). $Pam_3CS(K)4GC$ was dissolved to 20 mg/ml in DMSO and added to a final concentration of 0.6 mg/ml. After incubating at room temperature for 3 hours, the reaction tube was placed in a 37°C

bath for 10 minutes to cause TX-114 droplet formation. The sample was then centrifuged at 16,000xg for 10 minutes to separate the detergent and aqueous phases. After drawing off the aqueous phase the detergent phase was resuspended to the original reaction volume with Buffer A. The total, detergent, and aqueous samples were then analyzed by SDS-PAGE.

Protein characterization: Proteins-lipopeptide conjugates were characterized by SDS-PAGE. Samples(typically 5 μg) were diluted in SDS-PAGE sample buffer (0.1M Tris, pH 8.0/4% SDS/25% glycerol/0.1M DTT). The samples were boiled for 5 minutes and loaded onto a 10% SDS polyacrylamide gel. Following electrophoresis, gels were stained with coomassie blue to visualize protein bands.

Results and Discussion

Input HA1-1His_6(PR8)Bv (SEQ ID NO: 179) runs as a doublet band on SDS-PAGE. This has been seen for all of the HA1-1His_6 proteins made in *Baculovirus* to date and is likely due to differences in glycosylation. HA1-1His_6(PR8)Bv (SEQ ID NO: 179) which has been derivatized with Sulfo-SMCC but not conjugated to lipopeptide stays in the aqueous phase during TX-114 separation, the expected behavior for a soluble, globular protein. HA1-1His_6(PR8)Bv (SEQ ID NO: 179) conjugated to Pam_3Cys shows 3 or 4 bands of higher molecular weight than the derivatized input protein, consistent with covalent modification by the lipopeptide. On TX-114 phase separation the higher mw species segregate primarily into the detergent phase while the lower species (corresponding to the unmodified protein) stay in the aqueous phase. This behavior is consistent with the commonly observed segregation of lipoproteins into detergent droplets and confirms that HA1-1His_6(PR8)Bv (SEQ ID NO: 179) has been covalently modified with the $Pam_3CS(K)_4GC$ (SEQ ID NO: 335) lipopeptide. The higher molecular-weight bands of approximately 60, 120, and 180 kDa, which are seen in the SMCC-derivatized protein but not in the input, are covalent HA1-1 multimers formed by the crosslinker. This multimerization would be prevented, and the yield of singly lipidated HA1-1 improved, by engineering a single cysteine into HA1-1 and using a cysteine-specific rather than a lysine-specific crosslinker. Such a strategy would also

give better control over the site of attachment of the lipopeptide on HA1-1, giving a more homogeneous product and possibly improving immunogenicity

SEQ ID NO: 312, amino acid sequence of STF2

MAQVINTNSLSLLTQNNLNKSQSALGTAIERLSSGLRINSAKDDAAGQAIANRFTANIKG
LTQASRNANDGISIAQTTEGALNEINNNLQRVRELAVQSANSTNSQSDLDSIQAEITQRL
NEIDRVSGQTQFNGVKVLAQDNTLTIQVGANDGETIDIDLKQINSQTLGLDSLNVQKAYD
VKDTAVTTKAYANNGTTLDVSGLDDAAIKAATGGTNGTASVTGGAVKFDADNNKYFVTIG
GFTGADAAKNGDYEVNVATDGTVTLAAGATKTTMPAGATTKTEVQELKDTPAVVSADAKN
ALIAGGVDATDANGAELVKMSYTDKNGKTIEGGYALKAGDKYYAADYDEATGAIKAKTTS
YTAADGTTKTAANQLGGVDGKTEVVTIDGKTYNASKAAGHDFKAQPELAEAAAKTTENPL
QKIDAALAQVDALRSDLGAVQNRFNSAITNLGNTVNNLSEARSRIEDSDYATEVSNMSRA
QILQQAGTSVLAQANQVPQNVLSLLR

SEQ ID NO: 313, amino acid sequence of STF2Δ

MAQVINTNSLSLLTQNNLNKSQSALGTAIERLSSGLRINSAKDDAAGQAIANRFTANIKG
LTQASRNANDGISIAQTTEGALNEINNNLQRVRELAVQSANSTNSQSDLDSIQAEITQRL
NEIDRVSGQTQFNGVKVLAQDNTLTIQVGANDGETIDIDLKQINSQTLGLDSLNVHGAPV
DPASPWTENPLQKIDAALAQVDALRSDLGAVQNRFNSAITNLGNTVNNLSEARSRIEDSD
YATEVSNMSRAQILQQAGTSVLAQANQVPQNVLSLLR

SEQ ID NO: 314, nucleotide sequence of STF2Δ

ATGGCACAAGTAATCAACACTAACAGTCTGTCGCTGCTGACCCAGAATAACCTGAACAAA
TCCCAGTCCGCACTGGGCACCGCTATCGAGCGTCTGTCTTCTGGTCTGCGTATCAACAGC
GCGAAAGACGATGCGGCAGGTCAGGCGATTGCTAACCGTTTCACCGCGAACATCAAAGGT
CTGACTCAGGCTTCCCGTAACGCTAACGACGGTATCTCCATTGCGCAGACCACTGAAGGC
GCGCTGAACGAAATCAACAACAACCTGCAGCGTGTGCGTGAACTGGCGGTTCAGTCTGCT
AACAGCACCAACTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATCACCCAGCGCCTG
AACGAAATCGACCGTGTATCCGGCCAGACTCAGTTCAACGGCGTGAAAGTCCTGGCGCAG
GACAACACCCTGACCATCCAGGTTGGCGCCAACGACGGTGAAACTATCGATATCGATCTG
AAGCAGATCAACTCTCAGACCCTGGGTCTGGACTCACTGAACGTGCATGGAGCGCCGGTG
GATCCTGCTAGCCCATGGACCGAAAACCCGCTGCAGAAAATTGATGCCGCGCTGGCGCAG
GTGGATGCGCTGCGCTCTGATCTGGGTGCGGTACAAAACCGTTTCAACTCTGCTATCACC
AACCTGGGCAATACCGTAAACAATCTGTCTGAAGCGCGTAGCCGTATCGAAGATTCCGAC
TACGCGACCGAAGTTTCCAACATGTCTCGCGCGCAGATTTTGCAGCAGGCCGGTACTTCC
GTTCTGGCGCAGGCTAACCAGGTCCCGCAGAACGTGCTGTCTCTGTTACGT

SEQ ID NO: 315, nucleotide sequence of STF2

ATGGCACAAGTAATCAACACTAACAGTCTGTCGCTGCTGACCCAGAATAACCTGAACAAA
TCCCAGTCCGCACTGGGCACCGCTATCGAGCGTCTGTCTTCTGGTCTGCGTATCAACAGC
GCGAAAGACGATGCGGCAGGTCAGGCGATTGCTAACCGTTTCACCGCGAACATCAAAGGT
CTGACTCAGGCTTCCCGTAACGCTAACGACGGTATCTCCATTGCGCAGACCACTGAAGGC
GCGCTGAACGAAATCAACAACAACCTGCAGCGTGTGCGTGAACTGGCGGTTCAGTCTGCT
AACAGCACCAACTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATCACCCAGCGCCTG
AACGAAATCGACCGTGTATCCGGCCAGACTCAGTTCAACGGCGTGAAAGTCCTGGCGCAG
GACAACACCCTGACCATCCAGGTTGGCGCCAACGACGGTGAAACTATCGATATCGATCTG
AAGCAGATCAACTCTCAGACCCTGGGTCTGGACTCACTGAACGTGCAGAAAGCGTATGAT
GTGAAAGATACAGCAGTAACAACGAAAGCTTATGCCAATAATGGTACTACACTGGATGTA
TCGGGTCTTGATGATGCAGCTATTAAAGCGGCTACGGGTGGTACGAATGGTACGGCTTCT
GTAACCGGTGGTGCGGTTAAATTTGACGCAGATAATAACAAGTACTTTGTTACTATTGGT
GGCTTTACTGGTGCTGATGCCGCCAAAAATGGCGATTATGAAGTTAACGTTGCTACTGAC
GGTACAGTAACCCTTGCGGCTGGCGCAACTAAAACCACAATGCCTGCTGGTGCGACAACT

AAAACAGAAGTACAGGAGTTAAAAGATACACCGGCAGTTGTTTCAGCAGATGCTAAAAAT
GCCTTAATTGCTGGCGGCGTTGACGCTACCGATGCTAATGGCGCTGAGTTGGTCAAAATG
TCTTATACCGATAAAAATGGTAAGACAATTGAAGGCGGTTATGCGCTTAAAGCTGGCGAT
AAGTATTACGCCGCAGATTACGATGAAGCGACAGGAGCAATTAAAGCTAAAACTACAAGT
TATACTGCTGCTGACGGCACTACCAAAACAGCGGCTAACCAACTGGGTGGCGTAGACGGT
AAAACCGAAGTCGTTACTATCGACGGTAAAACCTACAATGCCAGCAAAGCCGCTGGTCAT
GATTTCAAAGCACAACCAGAGCTGGCGGAAGCAGCCGCTAAAACCACCGAAAACCCGCTG
CAGAAAATTGATGCCGCGCTGGCGCAGGTGGATGCGCTGCGCTCTGATCTGGGTGCGGTA
CAAAACCGTTTCAACTCTGCTATCACCAACCTGGGCAATACCGTAAACAATCTGTCTGAA
GCGCGTAGCCGTATCGAAGATTCCGACTACGCGACCGAAGTTTCCAACATGTCTCGCGCG
CAGATTCTGCAGCAGGCCGGTACTTCCGTTCTGGCGCAGGCTAACCAGGTCCCGCAGAAC
GTGCTGTCTCTGTTACGT

SEQ ID NO: 316, nucleotide sequence of STF2.OVA
ATGGCACAAGTAATCAACACTAACAGTCTGTCGCTGCTGACCCAGAATAACCTGAACAAA
TCCCAGTCCGCACTGGGCACCGCTATCGAGCGTCTGTCTTCTGGTCTGCGTATCAACAGC
GCGAAAGACGATGCGGCAGGTCAGGCGATTGCTAACCGTTTCACCGCGAACATCAAAGGT
CTGACTCAGGCTTCCCGTAACGCTAACGACGGTATCTCCATTGCGCAGACCACTGAAGGC
GCGCTGAACGAAATCAACAACAACCTGCAGCGTGTGCGTGAACTGGCGGTTCAGTCTGCT
AACAGCACCAACTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATCACCCAGCGCCTG
AACGAAATCGACCGTGTATCCGGCCAGACTCAGTTCAACGGCGTGAAAGTCCTGGCGCAG
GACAACACCCTGACCATCCAGGTTGGCGCCAACGACGGTGAAACTATCGATATCGATCTG
AAGCAGATCAACTCTCAGACCCTGGGTCTGGACTCACTGAACGTGCAGAAAGCGTATGAT
GTGAAAGATACAGCAGTAACAACGAAAGCTTATGCCAATAATGGTACTACACTGGATGTA
TCGGGTCTTGATGATGCAGCTATTAAAGCGGCTACGGGTGGTACGAATGGTACGGCTTCT
GTAACCGGTGGTGCGGTTAAATTTGACGCAGATAATAACAAGTACTTTGTTACTATTGGT
GGCTTTACTGGTGCTGATGCCGCCAAAAATGGCGATTATGAAGTTAACGTTGCTACTGAC
GGTACAGTAACCCTTGCGGCTGGCGCAACTAAAACCACAATGCCTGCTGGTGCGACAACT
AAAACAGAAGTACAGGAGTTAAAAGATACACCGGCAGTTGTTTCAGCAGATGCTAAAAAT
GCCTTAATTGCTGGCGGCGTTGACGCTACCGATGCTAATGGCGCTGAGTTGGTCAAAATG
TCTTATACCGATAAAAATGGTAAGACAATTGAAGGCGGTTATGCGCTTAAAGCTGGCGAT
AAGTATTACGCCGCAGATTACGATGAAGCGACAGGAGCAATTAAAGCTAAAACCACAAGT
TATACTGCTGCTGACGGCACTACCAAAACAGCGGCTAACCAACTGGGTGGCGTAGACGGT
AAAACCGAAGTCGTTACTATCGACGGTAAAACCTACAATGCCAGCAAAGCCGCTGGTCAT
GATTTCAAAGCACAACCAGAGCTGGCGGAAGCAGCCGCTAAAACCACCGAAAACCCGCTG
CAGAAAATTGATGCCGCGCTGGCGCAGGTGGATGCGCTGCGCTCTGATCTGGGTGCGGTA
CAAAACCGTTTCAACTCTGCTATCACCAACCTGGGCAATACCGTAAACAATCTGTCTGAA
GCGCGTAGCCGTATCGAAGATTCCGACTACGCGACCGAAGTTTCCAACATGTCTCGCGCG
CAGATTTTGCAGCAGGCCGGTACTTCCGTTCTGGCGCAGGCTAACCAGGTCCCGCAGAAC
GTGCTGTCTCTGTTACGTCTCGAGGGCTCCATCGGCGCAGCAAGCATGGAATTTTGTTTT
GATGTATTCAAGGAGCTCAAAGTCCACCATGCCAATGAGAACATCTTCTACTGCCCCATT
GCCATCATGTCAGCTCTAGCCATGGTATACCTGGGTGCAAAAGACAGCACCAGGACACAA
ATAAATAAGGTTGTTCGCTTTGATAAACTTCCAGGATTCGGAGACAGTATTGAAGCTCAG
TGTGGCACATCTGTAAACGTTCACTCTTCACTTAGAGACATCCTCAACCAAATCACCAAA
CCAAATGATGTTTATTCGTTCAGCCTTGCCAGTAGACTTTATGCTGAAGAGAGATACCCA
ATCCTGCCAGAATACTTGCAGTGTGTGAAGGAACTGTATAGAGGAGGCTTGGAACCTATC
AACTTTCAAACAGCTGCAGATCAAGCCAGAGAGCTCATCAATTCCTGGGTAGAAAGTCAG
ACAAATGGAATTATCAGAAATGTCCTTCAGCCAAGCTCCGTGGATTCTCAAACTGCAATG
GTTCTGGTTAATGCCATTGTCTTCAAAGGACTGTGGGAGAAAGCATTTAAGGATGAAGAC
ACACAAGCAATGCCTTTCAGAGTGACTGAGCAAGAAAGCAAACCTGTGCAGATGATGTAC
CAGATTGGTTTATTTAGAGTGGCATCAATGGCTTCTGAGAAAATGAAGATCCTGGAGCTT
CCATTTGCCAGTGGGACAATGAGCATGTTGGTGCTGTTGCCTGATGAAGTCTCAGGCCTT
GAGCAGCTTGAGAGTATAATCAACTTTGAAAAACTGACTGAATGGACCAGTTCTAATGTT
ATGGAAGAGAGGAAGATCAAAGTGTACTTACCTCGCATGAAGATGGAGGAAAAATACAAC

CTCACATCTGTCTTAATGGCTATGGGCATTACTGACGTGTTTAGCTCTTCAGCCAATCTG
TCTGGCATCTCCTCAGCAGAGAGCCTGAAGATATCTCAAGCTGTCCATGCAGCACATGCA
GAAATCAATGAAGCAGGCAGAGAGGTGGTAGGGTCAGCAGAGGCTGGAGTGGATGCTGCA
AGCGTCTCTGAAGAATTTAGGGCTGACCATCCATTCCTCTTCTGTATCAAGCACATCGCA
ACCAACGCCGTTCTCTTCTTTGGCAGATGTGTTTCCCCTTCGAAGCTTGAAGGTAAGCCT
ATCCCTAACCCTCTCCTCGGTCTCGATTCTACGCGTACCGGTCATCATCACCATCACCAT
TGA

SEQ ID NO: 317, nucleotide sequence of STF28BGF-1
CTCGGGAGATCTGCACAAGTAATCAACACTAACAGTCT

SEQ ID NO: 318, nucleotide sequence of STF28MCR-1
CCATGGGCTAGCAGGATCCACCGGCGCTCCCTGCACGTTCA

SEQ ID NO: 319, nucleotide sequence of STF28MCF-2
GGAGCGCCGGTGGATCCTGCTAGCCCATGGACCGAAAACCCG

SEQ ID NO: 320, nucleotide sequence of STF28ECR-2
TCTGCAGAATTCACGTAACAGAGACAGCACGTTCTGCGGGACGTCCCGCAGAACGTGCTG
TCTCTGTTACGTGAATTCTGCAGA

SEQ ID NO: 312, nucleotide sequence of 3'Forward1
ACTGAGTGCATATGGCACAAGTAATCAACACTAACAG

SEQ ID NO: 322, nucleotide sequence of 3'Reverse1
GACTGACTGCTCAGCCTATTAGCAGAGCGGCCGCCACTGTGCTGGATATCTAGAG

SEQ ID NO: 323 nucleotide sequence of 5'Forward2
AGTCAGGCCATATGTGCGCACAAGTAATCAACACTAACAGTCTG

SEQ ID NO: 324, nucleotide sequence of 5'Reverse2
GACTGACTGCTCAGCCTATTAACGTAACAGAGACAGCACGTTCTGCGGGACCTGGTTAG

SEQ ID NO: 325, nucleotide sequence of STF2Δ.3'Cys
ATGGCACAAGTAATCAACACTAACAGTCTGTCGCTGCTGACCCAGAATAACCTGAACAAA
TCCCAGTCCGCACTGGGCACCGCTATCGAGCGTCTGTCTTCTGGTCTGCGTATCAACAGC
GCGAAAGACGATGCGGCAGGTCAGGCGATTGCTAACCGTTTCACCGCGAACATCAAAGGT
CTGACTCAGGCTTCCCGTAACGCTAACGACGGTATCTCCATTGCGCAGACCACTGAAGGC
GCGCTGAACGAAATCAACAACAACCTGCAGCGTGTGCGTGAACTGGCGGTTCAGTCTGCT
AACAGCACCAACTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATCACCCAGCGCCTG
AACGAAATCGACCGTGTATCCGGCCAGACTCAGTTCAACGGCGTGAAAGTCCTGGCGCAG
GACAACACCCTGACCATCCAGGTTGGCGCCAACGACGGTGAAACTATCGATATCGATCTG
AAGCAGATCAACTCTCAGACCCTGGGTCTGGACTCACTGAACGTGCATGGAGCGCCGGTG
GATCCTGCTAGCCCATGGACCGAAAACCCGCTGCAGAAAATTGATGCCGCGCTGGCGCAG
GTGGATGCGCTGCGCTCTGATCTGGGTGCGGTACAAAACCGTTTCAACTCTGCTATCACC
AACCTGGGCAATACCGTAAACAATCTGTCTGAAGCGCGTAGCCGTATCGAAGATTCCGAC
TACGCGACCGAAGTTTCCAACATGTCTCGCGCGCAGATTTTGCAGCAGGCCGGTACTTCC
GTTCTGGCGCAGGCTAACCAGGTCCCGCAGAACGTGCTGTCTCTGTTACGTGAATTCTCT
AGATATCCAGCACAGTGGCGGCCGCTCTGC

SEQ ID NO: 326, nucleotide sequence of 5'Cys.STF2Δ
ATGTGCGCACAAGTAATCAACACTAACAGTCTGTCGCTGCTGACCCAGAATAACCTGAAC
AAATCCCAGTCCGCACTGGGCACCGCTATCGAGCGTCTGTCTTCTGGTCTGCGTATCAAC
AGCGCGAAAGACGATGCGGCAGGTCAGGCGATTGCTAACCGTTTCACCGCGAACATCAAA

GGTCTGACTCAGGCTTCCCGTAACGCTAACGACGGTATCTCCATTGCGCAGACCACTGAA
GGCGCGCTGAACGAAATCAACAACAACCTGCAGCGTGTGCGTGAACTGGCGGTTCAGTCT
GCTAACAGCACCAACTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATCACCCAGCGC
CTGAACGAAATCGACCGTGTATCCGGCCAGACTCAGTTCAACGGCGTGAAAGTCCTGGCG
CAGGACAACACCCTGACCATCCAGGTTGGCGCCAACGACGGTGAAACTATCGATATCGAT
CTGAAGCAGATCAACTCTCAGACCCTGGGTCTGGACTCACTGAACGTGCATGGAGCGCCG
GTGGATCCTGCTAGCCCATGGACCGAAAACCCGCTGCAGAAAATTGATGCCGCGCTGGCG
CAGGTGGATGCGCTGCGCTCTGATCTGGGTGCGGTACAAAACCGTTTCAACTCTGCTATC
ACCAACCTGGGCAATACCGTAAACAATCTGTCTGAAGCGCGTAGCCGTATCGAAGATTCC
GACTACGCGACCGAAGTTTCCAACATGTCTCGCGCGCAGATTTTGCAGCAGGCCGGTACT
TCCGTTCTGGCGCAGGCTAACCAGGTCCCGCAGAACGTGCTGTCTCTGTTACGT

SEQ ID NO: 327, nucleotide sequence of HingeForward1
ACTGAGTGCATATGGCACAAGTAATCAACACTAACAG

SEQ ID NO: 328, nucleotide sequence of HingeReverse3
GGTCCATGGGCAAGCAGGATCCACCGGCGCT

SEQ ID NO: 329, nucleotide sequence of HingeForward2
AGCGCCGGTGGATCCTGCTTGCCCATGGACC

SEQ ID NO: 330, nucleotide sequence of HingeReverse4
GACTGACTGCTCAGCCTATTAACGTAACAGAGACAGCACGTTCTGCGGGACCTGGTTAG

SEQ ID NO: 331, nucleotide sequence of STF2Δ.HingeCys
ATGGCACAAGTAATCAACACTAACAGTCTGTCGCTGCTGACCCAGAATAACCTGAACAAA
TCCCAGTCCGCACTGGGCACCGCTATCGAGCGTCTGTCTTCTGGTCTGCGTATCAACAGC
GCGAAAGACGATGCGGCAGGTCAGGCGATTGCTAACCGTTTCACCGCGAACATCAAAGGT
CTGACTCAGGCTTCCCGTAACGCTAACGACGGTATCTCCATTGCGCAGACCACTGAAGGC
GCGCTGAACGAAATCAACAACAACCTGCAGCGTGTGCGTGAACTGGCGGTTCAGTCTGCT
AACAGCACCAACTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATCACCCAGCGCCTG
AACGAAATCGACCGTGTATCCGGCCAGACTCAGTTCAACGGCGTGAAAGTCCTGGCGCAG
GACAACACCCTGACCATCCAGGTTGGCGCCAACGACGGTGAAACTATCGATATCGATCTG
AAGCAGATCAACTCTCAGACCCTGGGTCTGGACTCACTGAACGTGCATGGAGCGCCGGTG
GATCCTGCTTGCCCATGGACCGAAAACCCGCTGCAGAAAATTGATGCCGCGCTGGCGCAG
GTGGATGCGCTGCGCTCTGATCTGGGTGCGGTACAAAACCGTTTCAACTCTGCTATCACC
AACCTGGGCAATACCGTAAACAATCTGTCTGAAGCGCGTAGCCGTATCGAAGATTCCGAC
TACGCGACCGAAGTTTCCAACATGTCTCGCGCGCAGATTTTGCAGCAGGCCGGTACTTCC
GTTCTGGCGCAGGCTAACCAGGTCCCGCAGAACGTGCTGTCTCTGTTACGT

SEQ ID NO: 332, amino acid sequence of STF2Δ.3'Cys
MAQVINTNSLSLLTQNNLNKSQSALGTAIERLSSGLRINSAKDDAAGQAIANRFTANIKG
LTQASRNANDGISIAQTTEGALNEINNNLQRVRELAVQSANSTNSQSDLDSIQAEITQRL
NEIDRVSGQTQFNGVKVLAQDNTLTIQVGANDGETIDIDLKQINSQTLGLDSLNVHGAPV
DPASPWTENPLQKIDAALAQVDALRSDLGAVQNRFNSAITNLGNTVNNLSEARSRIEDSD
YATEVSNMSRAQILQQAGTSVLAQANQVPQNVLSLLREFSRYPAQWRPL[C]

SEQ ID NO: 333, amino acid sequence of STF2Δ.HingeCys
MAQVINTNSLSLLTQNNLNKSQSALGTAIERLSSGLRINSAKDDAAGQAIANRFTANIKG
LTQASRNANDGISIAQTTEGALNEINNNLQRVRELAVQSANSTNSQSDLDSIQAEITQRL
NEIDRVSGQTQFNGVKVLAQDNTLTIQVGANDGETIDIDLKQINSQTLGLDSLNVHGAPV
DPA[C]PWTENPLQKIDAALAQVDALRSDLGAVQNRFNSAITNLGNTVNNLSEARSRIEDSD
YATEVSNMSRAQILQQAGTSVLAQANQVPQNVLSLLR

SEQ ID NO: 334, amino acid sequence of 5'Cys.STF2Δ
MCAQVINTNSLSLLTQNNLNKSQSALGTAIERLSSGLRINSAKDDAAGQAIANRFTANIK
GLTQASRNANDGISIAQTTEGALNEINNNLQRVRELAVQSANSTNSQSDLDSIQAEITQR
LNEIDRVSGQTQFNGVKVLAQDNTLTIQVGANDGETIDIDLKQINSQTLGLDSLNVHGAP
VDPASPWTENPLQKIDAALAQVDALRSDLGAVQNRFNSAITNLGNTVNNLSEARSRIEDS
DYATEVSNMSRAQILQQAGTSVLAQANQVPQNVLSLLR

SEQ ID NO: 335,
CSKKKKGC

SEQ ID NO: 336,
GAPVDPASPW

EXAMPLE 14 CLONING AND EXPRESSION OF STF2.4xH1C1, A FUSION PROTEIN COMPRISING FLAGELLIN (TLR5 AGONIST) AND THE MATURATIONAL CLEAVAGE SITE PEPTIDE OF INFLUENZA A HEMAGGLUTININ

Materials and Methods

Construct design: To facilitate cloning of target genes in fusion with flagellin, a cassette plasmid containing a unique *Blp*I site at the 3' end of the flagellin gene was used (STF2.blp, SEQ ID NO: 340). This cassette was made by introducing a silent mutation (5'-GTGCTGAGCCTGTTACGT-3') at nucleotides 1501 to 1518 of STF2, creating the unique *Blp*I site in the plasmid cassette, pET24/STF2.blp (SEQ ID NO: 340). Synthetic genes encoding one, two, three and four copies of the maturational cleavage site fragment (H1C1) of influenza A subtype H1 (NIPSIQSRGLFGAIAGFIE; SEQ ID NO: 341)were codon-optimized for *E. coli* expression and obtained from a commercial vendor (DNA2.0 Inc., Menlo Park, CA). The synthetic genes were excised with *Blp*I enzyme and ligated by compatible ends to pET24/STF2.blp which had been treated with *Blp*I and bacterial alkaline phosphatase (BAP). These constructs were designated as STF2.1xH1C1, STF2.2xH1C1, STF2.3xH1C1 and STF2.4xH1C1 respectively (SEQ ID NO: 342; SEQ ID NO: 343; SEQ ID NO: 344; SEQ ID NO: 345).

Similarly, a fusion gene constructed to contain the H1C1 sequence and flanked by a pair of cysteines to facilitate loop formation of the concatemers was designated as STF2.4xH1C2 (SEQ ID NO: 346). A similar approach was employed to generate a construct harboring four tandem copies of the cleavage fragment of influenza A subtype H5 (RERRRKKRGLFGAIAGFIE; SEQ ID NO: 347)

designated as construct STF2.4xH5C1 (SEQ ID NO: 348). The consensus cleavage fragments representing subtypes H3 (NVPEKQTRGIFGAIAGFIE; SEQ ID NO: 349), H2 (NVPQIESRGLFGAIAGFIE; SEQ ID NO: 350) and the B-strain, B/HA (PAKLLKERGFFGAIAGFLE; SEQ ID NO: 351) are thus amenable to this experimental approach as outlined above. In each case, the constructed plasmids were used to transform competent *E. coli* TOP10 cells and putative recombinants were identified by PCR screening and restriction mapping analysis.

The integrity of the constructs was verified by DNA sequencing and they were used to transform the expression host, BLR3 (DE3) (Novagen, San Diego, CA; Cat #69053). Transformants were selected on plates containing kanamycin (50 μg/mL), tetracycline (5 μg/mL) and glucose (0.5%). Colonies were picked and inoculated into 2 ml of LB medium supplemented with 25 μg/ml kanamycin, 12.5 μg/ml tetracycline and 0.5% glucose and grown overnight. Aliquots of these cultures were used to inoculate fresh cultures in the same medium formulation, which were cultured until they reached an OD_{600} = 0.6, at which time protein expression was induced by the addition of 1 mM IPTG and culturing for 3 hours at 37°C. The cells were harvested and analyzed for protein expression.

SDS-PAGE and Western blot: Protein expression and identity were determined by gel electrophoresis and immunoblot analysis. Cells were harvested by centrifugation and lysed in Laemmli buffer. An aliquot of 10 μl of each lysate was diluted in SDS-PAGE sample buffer with or without 100 mM DTT as a reductant. The samples were boiled for 5 minutes and loaded onto a 10% SDS polyacrylamide gel and electrophoresed (SDS-PAGE). The gel was stained with Coomassie R-250 (Bio-Rad; Hercules, CA) to visualize protein bands. For western blot, 0.5 μl/lane cell lysate was electrophoresed and electrotransfered to a PVDF membrane and blocked with 5% (w/v) dry milk . The membrane was probed with anti-flagellin antibody 6H11 (Inotek; Beverly, MA) After probing with alkaline phosophatase-conjugated secondary antibodies (Pierce; Rockland, IL), protein bands were visualized with an alkaline phosphatase chromogenic substrate (Promega, Madison, WI). Bacterial clones which yielded protein bands of the correct molecular weight and reactive with the appropriate antibodies were selected for production of protein for use in biological assays.

Results and Discussion

Cloning and expression of cleavage site constructs: The maturational cleavage site fragment of hemagglutinin precursor is well-conserved across subtypes of influenza strains A and B, and could therefore be a target for the development of a universal vaccine effective against most circulating strains of influenza. A series of plasmids encoding a fusion of the TLR5 ligand with cleavage fragment was generated and expressed in *E. coli* strain BLR(DE3). As assayed by Coomassie blue staining of the SDS-PAGE gel and confirmed by immunoblot assays, the *E. coli* strains harboring the constructs STF2.1xH1C1, STF2.2xH1C1, STF2.3xH1C1, and STF2.4xH1C1 displayed bands that correspond to the predicted molecular weights of 55, 58, 60 and 62 kDa respectively .. Similarly, BLR(DE3) strains expressing the constructs STF2.4xH1C2 and STF2.4xH5C1 display recombinant fusion proteins migrating with the apparent molecular weight of 62 Kda and 66 Kda respectively. These data indicate that a fusion of flagellin and cleavage fragment of HA is abundantly expressed in *E coli*.

EXPRESSION AND PURIFICATION OF STF2.4xH1C1, A FUSION PROTEIN COMPRISING FLAGELLIN (TLR5 AGONIST) AND THE MATURATIONAL CLEAVAGE SITE PEPTIDE OF INFLUENZA A HEMAGGLUTININ

Materials and methods

Bacterial cell growth and cell lysis: The STF2.4xH1C1 (SEQ ID NO: 345) construct was expressed in the *E. coli* host strain BLR(DE3). The strain was retrieved from a glycerol stock and grown in shake flasks to a final volume of 12 L. Cells were grown in LB medium containing 50 µg/ml kanamycin, 12.5 µg/ml tetracycline, 0.5% dextrose to OD_{600} = 0.6 and induced with 1 mM IPTG for 3 h at 37°C. The cells were harvested by centrifugation (7000 rpm x 7 minutes in a Sorvall RC5C centrifuge) and resuspended in 20 mM Tris-HCl, pH 8.0, 1 µg/ml DNAseI, 1 mM PMSF, protease inhibitor cocktail and 1 mg/ml lysozyme. The cells were then lysed by two passes through a microfluidizer (Microfluidics; Newton, MA) at 15,000 psi. The lysate was centrifuged at 45,000 g for one hour in a Beckman Optima L ultracentrifuge (Beckman Coulter; Fullerton, CA) to separate the soluble and insoluble fractions.

Purification of STF2.4xH1C1 (SEQ ID NO: 345): After centrifugation, the supernatant fraction was collected and passed through a Q sepharose Fast Flow column (GE/Amersham Biosciences; Piscataway, NJ). The flow-through fraction from this step was supplemented with Triton X-100 (Sigma; St. Louis, MO) to a final concentration of 1% (w/v) and passed through the same Q sepharose column. The flow through fraction was collected again and supplemented with urea to a final concentration of 8 M and citric acid to a final concentration of 20 mM. After adjusting the pH to 3.5 with concentrated HCl, the solution was passed over a Source S column (GE/Amersham Biosciences; Piscataway, NJ) equilibrated with 20 mM citric acid, pH 3.5.

The column was then washed with 10 column volumes of equilibration buffer supplemented with 1% (w/v) Triton X-100 to remove endotoxin. The protein was then eluted in a 5-column volume linear gradient of 0 to 1 M NaCl in equilibration buffer. The STF2.4xH1C1 was then re-folded by rapid dilution to a final concentration of 0.1 mg/ml protein in refolding buffer [0.1 M Tris-HCl, pH 8.0, 0.1 M NaC, 1% (w/v) glycerol]. The refolded protein was concentrated using a pressurized ultrafiltration stir-cell (Millipore; Billerica, MA) and fractionated on a Superdex 200 size-exclusion column (GE/Amersham Biosciences; Piscataway, NJ).

SDS-PAGE and Western Blot analysis: Protein identity was determined, and purity estimated, by SDS-PAGE. An aliquot of 5 µg of each sample was diluted in SDS-PAGE sample buffer with or without 100 mM DTT as a reductant. The samples were boiled for 5 minutes and loaded onto a 10% SDS polyacrylamide gel (LifeGels; French's Forrest, New South Wales, AUS) and electrophoresed. The gel was stained with Coomassie R-250 (Bio-Rad; Hercules, CA) to visualize protein bands. For western blot, 0.5 µg/lane total protein was electrophoresed as described above and then electro-transfered to a PVDF membrane and blocked with 5% (w/v) dry milk before probing with anti-flagellin antibody (Inotek; Beverly, MA) or serum from mice immunized with a synthetic, lipidated H1C1 peptide. After probing with alkaline phosophatase-conjugated secondary antibodies (Pierce; Rockland, IL), protein bands were visualized with an alkaline phosphatase chromogenic substrate (Promega; Madison, WI).

Protein assay: Total protein concentration for all proteins was determined using the Micro BCA (bicinchoninic acid) Assay (Pierce; Rockford IL) in the microplate format, using bovine serum albumin as a standard, according to the manufacturer's instructions.

Endotoxin assay: Endotoxin levels for all proteins were determined using the QCL-1000 Quantitative Chromogenic LAL test kit (Cambrex; E. Rutherford, NJ) , following the manufacturer's instructions for the microplate method.

TLR5 bioactivity assay: HEK293 cells constitutively express TLR5, and secrete several soluble factors, including IL-8, in response to TLR5 signaling. Cells were seeded in 96-well microplates (50,000 cells/well), and the STF2.4xH1C1 test protein was added. The next day, the conditioned medium was harvested, transferred to a clean 96-well microplate, and frozen at -20°C. After thawing, the conditioned medium was assayed for the presence of IL-8 in a sandwich ELISA using an anti-human IL-8 matched antibody pair (Pierce, Rockland, IL; #M801E and #M802B) following the manufacturer's instructions. Optical density was measured using a microplate spectrophotometer (FARCyte, GE/Amersham; Piscataway, NJ).

Results and Discussion

The purified and refolded STF2.4xH1C1 protein (SEQ ID NO: 345) was found to be substantially aggregated as the majority of the protein fractionated in the void volume of the Superdex S200 column (Figure 26). The protein also had poor *in vitro* TLR5 activity, with an EC_{50} value approximately two orders of magnitude higher than the standard flagellin fusion protein STF2.OVA. The purified protein reacted with serum from mice immunized with a synthetic, lipidated H1C1 peptide.

EXPRESSION AND PURIFICATION OF STF2.1xH1C1, A FUSION PROTEIN COMPRISING FLAGELLIN (TLR5 AGONIST) AND THE MATURATIONAL CLEAVAGE SITE PEPTIDE OF INFLUENZA A HEMAGGLUTININ

Materials and methods

Bacterial cell growth and cell lysis: The STF2.1xH1C1 (SEQ ID NO: 355) construct was expressed in the *E. coli* host strain BLR(DE3). The strain was retrieved from a glycerol stock and grown in shake flasks to a final volume of 12L.

Cells were grown in LB medium containing 50 μg/ml kanamycin/ 12.5 μg/ml tetracycline/0.5% dextrose to OD_{600} = 0.6 and induced with 1 mM IPTG for 3 hours at 37°C. The cells were harvested by centrifugation (7000 rpm x 7 minutes in a Sorvall RC5C centrifuge) and resuspended in 20 mM Tris-HCl, pH 8.0, 1 μg/ml DNAseI, 1 mM PMSF, protease inhibitor cocktail and 1 mg/ml lysozyme. The cells were then lysed by two passes through a microfluidizer (Microfluidics; Newton, MA) at 15,000 psi. The lysate was centrifuged at 45,000 g for one hour in a Beckman Optima L ultracentrifuge (Beckman Coulter; Fullerton, CA) to separate the soluble and insoluble fractions.

Purification of STF2.1xH1C1 (SEQ ID NO: 342): After centrifugation, the supernatant fraction was collected and passed through a Q sepharose Fast Flow column (GE/Amersham Biosciences; Piscataway, NJ). The flow-through fraction from this step was supplemented with Triton X-100 (Sigma; St. Louis, MO) to a final concentration of 1% (w/v) and passed through the same Q sepharose column. The flow-through fraction was collected again and supplemented with urea to a final concentration of 8 M and citric acid to a final concentration of 20 mM. After adjusting the pH to 3.5 with concentrated HCl, the solution was passed over a Source S column (GE/Amersham Biosciences; Piscataway, NJ) equilibrated with 20 mM citric acid, pH 3.5.

The column was then washed with 10 column volumes of equilibration buffer supplemented with 1% (w/v) Triton X-100 to remove endotoxin. The protein was then eluted in a 5-column volume linear gradient of 0 to 1 M NaCl in equilibration buffer. The STF2.1xH1C1 was re-folded by rapid dilution to a final concentration of 0.1 mg/ml protein in refolding buffer [0.1M Tris-HCl, pH 8.0/ 0.1M NaCl/ 1% (w/v) glycerol]. The refolded protein was concentrated using a pressurized ultrafiltration stir-cell (Millipore; Billerica, MA) and fractionated on a Superdex 200 size-exclusion column (GE/Amersham Biosciences; Piscataway, NJ).

SDS-PAGE and western blot analysis: Protein identity was determined, and purity estimated, by SDS-PAGE. An aliquot of 5 μg of each sample was diluted in SDS-PAGE sample buffer with or without 100 mM DTT as a reductant. The samples were boiled for 5 minutes and loaded onto a 10% SDS polyacrylamide gel (LifeGels; French's Forrest, New South Wales, AUS) and electrophoresed. The gel

was stained with Coomassie R-250 (Bio-Rad; Hercules, CA) to visualize protein bands. For western blot, 0.5 μg/lane total protein was electrophoresed as described above and the gels were then electro-transfered to a PVDF membrane and blocked with 5% (w/v) dry milk before probing with anti-flagellin antibody (Inotek; Beverly, MA) or serum from mice immunized with a Pam3Cys.H1C1 peptide (SEQ ID NO: 358). After probing with alkaline phosophatase-conjugated secondary antibodies (Pierce; Rockland, IL), protein bands were visualized with an alkaline phosphatase chromogenic substrate (Promega; Madison, WI).

Results and Discussion

The purified and refolded STF2.1xH1C1 protein (SEQ ID NO: 342) was found to be monomeric as judged by the elution profile on a Superdex 200 gel filtration column. The majority of the protein was found in the included volume with the major peak eluting at approximately 14 mls. This corresponds very closely with the known elution profile of purified, monomeric flagellin on this column. Almost no protein was seen eluting in the void volume, known to be approximately 7 mls for this column, demonstrating that virtually no aggregates are present.

IMMUNOGENICITY AND EFFICACY OF PAM3CYS.H1C1, A FUSION PEPTIDE COMPRISING PAM3CYS (TLR2 AGONIST) AND THE MATURATIONAL CLEAVAGE SITE PEPTIDE OF INFLUENZA A HEMAGGLUTININ

Materials and methods

Peptide design and synthesis: Pam3 (tri-palmoytyl) is the natural ligand for Toll-like Receptor 2 (TLR2), and it is natively expressed as the lipidation motif of bacterial lipoprotein (BLP, SEQ ID NO: 357). Pam3 can be made by chemical synthesis and conjugated to macromolecules such as proteins or coupled to the N-terminus of synthetic peptides using standard peptide synthesis chemistry. This strategy usually involves the synthesis of a peptide of interest concluded by the coupling of Pam3-modified cysteine (Pam3Cys) to the amino terminus to yield the lipopeptide of interest. This approach was used to synthesize a lipidated HA

cleavage fragment peptide, Pam3Cys.H1C1 (Pam3Cys-SLWSEENIPSIQSRGLFGAIAGFIEE, SEQ ID NO: 358).

Mice and immunization: Female BALB/c mice (Jackson Laboratory, Bar Harbor, ME) were used at the age of 6-8 weeks. Mice were divided into groups of 10 and received inguinal subcutaneous (s.c) immunizations on days 0 and 14 as follows:

8) PBS (phosphate buffered saline).

9) 20 μg of H1C1 native peptide (SEQ ID NO: 358) in saline buffer (10mM Histidine, 10mM Tris, 75mM NaCl, 5% (vol/vol) sucrose, 0.02% (w/v) Polysorbate-80, 0.1mM EDTA, 0.5% (v/v) ethanol, pH 7.2)

10) 20 μg of Pam3Cys.H1C1 peptide (SEQ ID NO: 358) in saline buffer

An additional group of five mice received an experimentally determined sublethal challenge with $8x10^1$ egg infectious dosages (EID) PR/8/34 and were allowed to convalesce for >21days. These animals were then used as immune convalescent positive controls during the challenge studies. Mice were bled on days 10 (primary) and 21 (boost), and sera were clarified by clotting and centrifugation and stored at -20°C.

Serum antibody determination: H1C1-specific IgG levels were determined by ELISA. 96-well ELISA plates (Costar (Cat #9018) Corning, NY) were coated overnight at 4°C with 100 μl/well H1C1 peptide (SEQ ID NO: 358) in PBS (5 μg/ml). Plates were blocked with 200 μl/well of Assay Diluent Buffer (ADB; BD Pharmingen, (Cat#: 555213) (San Diego, CA) for one hour at room temperature. The plates were washed three times in PBS+ 0.05% (v/v) Tween 20 (PBS-T). Dilutions of the sera in ADB were added (100 μl/well) and the plates were incubated overnight at 4°C. The plates were washed three times with PBS-T. HRP-labeled goat anti-mouse IgG antibodies (Jackson Immunochemical, West Grove, PA (Cat#: 115-035-146)) diluted in ADB were added (100 μl/well) and the plates were incubated at room temperature for 1 hour. The plates were washed three times with PBS-T. After adding TMB Ultra substrate (Pierce (Cat 34028), Rockford, IL)) and monitoring color development, A_{450} was measured on a Tecan Farcyte (Durham, NC) microplate spectrophotometer.

Influenza virus challenge of mice. To assess efficacy, mice immunized as described above were challenged on day 28 by intranasal administration of an LD_{90} (dose lethal to 90% of mice) ($8x10^3$ EID) of influenza A isolate PR/8/34. Animals were monitored daily for 21 days following the challenge for survival, weight loss and clinical presentation. The % weight loss was calculated based on the mean of ((Daily weight (g) /Initial weight (g) day 28) x 100) of each individual animal per group. Clinical scores were assigned as follows: 4 pts=healthy, 3 pts= reduced grooming, 2 pts= reduced physical activity and 1 pt = moribund. (Experimental results for clinical scores and weight loss reflect the results based on surviving animals on the day evaluated).

Results and Discussion

Induction of H1C1-specific antibody responses following immunization with Pam3Cys.H1C1: Mice were immunized with native H1C1 peptide (SEQ ID NO: 358) and the same peptide modified by the addition of an amino terminal Pam3Cys residue, to test the hypothesis that linkage of the peptide to a TLR2 ligand would increase its immunogenicity. Immunogenicity was determined by measuring levels of antibodies to native H1C1 (SEQ ID NO: 358) in the sera of the immunized mice. The results show clearly that mice immunized with the Pam3Cys-modified peptide generated higher antibody titers to the immunizing peptide backbone sequence than did mice immunized with the native peptide (SEQ ID NO: 358) (Figure 27).

Protection from lethal influenza virus challenge following immunization with Pam3Cys.H1C1 The results in Figure 27 demonstrated that immunization of mice with Pam3Cys.H1C1 generated an antibody response that recognized native H1C1 peptide. In order to evaluate efficacy, the same mice were challenged on day 28 with an LD_{90} ($8x10^3$EID) of PR/8/34 virus administered intra-nasally. Mice were monitored daily for 21 days following the challenge for survival, weight loss and clinical presentation. As shown in Figure 28, PBS-immunized mice died between 6 and 10 days post-challenge, while convalescent mice survived beyond the 21-day observation period. Mice immunized with native H1C1 (SEQ ID NO: 358) began to die by day 6 post-infection at a rate similar to the PBS control mice, although 2 of the 10 mice survived beyond the 21-day observation period. In contrast, and similar

to the convalescent mice, all mice immunized with Pam3Cys.H1C1 survived beyond the 21-day observation period. Thus, while H1C1 alone is not potently immunogenic (Figure 27) nor protective (Figure 28), the same sequence coupled to a TLR2 ligand is both immunogenic and protective.

SEQ ID NO: 340 STF2.blp

ATGGCACAAGTAATCAACACTAACAGTCTGTCGCTGCTGACCCAGAATAACCTGAACAAA
TCCCAGTCCGCACTGGGCACCGCTATCGAGCGTCTGTCTTCTGGTCTGCGTATCAACAGC
GCGAAAGACGATGCGGCAGGTCAGGCGATTGCTAACCGTTTCACCGCGAACATCAAAGGT
CTGACTCAGGCTTCCCGTAACGCTAACGACGGTATCTCCATTGCGCAGACCACTGAAGGC
GCGCTGAACGAAATCAACAACAACCTGCAGCGTGTGCGTGAACTGGCGGTTCAGTCTGCT
AACAGCACCAACTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATCACCCAGCGCCTG
AACGAAATCGACCGTGTATCCGGCCAGACTCAGTTCAACGGCGTGAAAGTCCTGGCGCAG
GACAACACCCTGACCATCCAGGTTGGCGCCAACGACGGTGAAACTATCGATATCGATCTG
AAGCAGATCAACTCTCAGACCCTGGGTCTGGACTCACTGAACGTGCAGAAAGCGTATGAT
GTGAAAGATACAGCAGTAACAACGAAAGCTTATGCCAATAATGGTACTACACTGGATGTA
TCGGGTCTTGATGATGCAGCTATTAAAGCGGCTACGGGTGGTACGAATGGTACGGCTTCT
GTAACCGGTGGTGCGGTTAAATTTGACGCAGATAATAACAAGTACTTTGTTACTATTGGT
GGCTTTACTGGTGCTGATGCCGCCAAAAATGGCGATTATGAAGTTAACGTTGCTACTGAC
GGTACAGTAACCCTTGCGGCTGGCGCAACTAAAACCACAATGCCTGCTGGTGCGACAACT
AAAACAGAAGTACAGGAGTTAAAAGATACACCGGCAGTTGTTTCAGCAGATGCTAAAAAT
GCCTTAATTGCTGGCGGCGTTGACGCTACCGATGCTAATGGCGCTGAGTTGGTCAAAATG
TCTTATACCGATAAAAATGGTAAGACAATTGAAGGCGGTTATGCGCTTAAAGCTGGCGAT
AAGTATTACGCCGCAGATTACGATGAAGCGACAGGAGCAATTAAAGCTAAAACTACAAGT
TATACTGCTGCTGACGGCACTACCAAAACAGCGGCTAACCAACTGGGTGGCGTAGACGGT
AAAACCGAAGTCGTTACTATCGACGGTAAAACCTACAATGCCAGCAAAGCCGCTGGTCAT
GATTTCAAAGCACAACCAGAGCTGGCGGAAGCAGCCGCTAAAACCACCGAAAACCCGCTG
CAGAAAATTGATGCCGCGCTGGCGCAGGTGGATGCGCTGCGCTCTGATCTGGGTGCGGTA
CAAAACCGTTTCAACTCTGCTATCACCAACCTGGGCAATACCGTAAACAATCTGTCTGAA
GCGCGTAGCCGTATCGAAGATTCCGACTACGCGACCGAAGTTTCCAACATGTCTCGCGCG
CAGATTCTGCAGCAGGCCGGTACTTCCGTTCTGGCGCAGGCTAACCAGGTCCCGCAGAAC
GTGCTGAGCCTGTTACGT

SEQ ID NO: 341 HA maturation cleavage fragment (H1N1)

NIPSIQSRGLFGAIAGFIE

SEQ ID NO: 342 STF2.1xH1C1

MAQVINTNSLSLLTQNNLNKSQSALGTAIERLSSGLRINSAKDDAAGQAIANRFTANIKG
LTQASRNANDGISIAQTTEGALNEINNNLQRVRELAVQSANSTNSQSDLDSIQAEITQRL
NEIDRVSGQTQFNGVKVLAQDNTLTIQVGANDGETIDIDLKQINSQTLGLDSLNVQKAYD
VKDTAVTTKAYANNGTTLDVSGLDDAAIKAATGGTNGTASVTGGAVKFDADNNKYFVTIG
GFTGADAAKNGDYEVNVATDGTVTLAAGATKTTMPAGATTKTEVQELKDTPAVVSADAKN
ALIAGGVDATDANGAELVKMSYTDKNGKTIEGGYALKAGDKYYAADYDEATGAIKAKTTS
YTAADGTTKTAANQLGGVDGKTEVVTIDGKTYNASKAAGHDFKAQPELAEAAAKTTENPL
QKIDAALAQVDALRSDLGAVQNRFNSAITNLGNTVNNLSEARSRIEDSDYATEVSNMSRA
QILQQAGTSVLAQANQVPQNVLSLLA**MEWENIPSIQSRGLFGAIAGFIE**

SEQ ID NO: 343 STF2.2xH1C1

MAQVINTNSLSLLTQNNLNKSQSALGTAIERLSSGLRINSAKDDAAGQAIANRFTANIKG
LTQASRNANDGISIAQTTEGALNEINNNLQRVRELAVQSANSTNSQSDLDSIQAEITQRL
NEIDRVSGQTQFNGVKVLAQDNTLTIQVGANDGETIDIDLKQINSQTLGLDSLNVQKAYD
VKDTAVTTKAYANNGTTLDVSGLDDAAIKAATGGTNGTASVTGGAVKFDADNNKYFVTIG
GFTGADAAKNGDYEVNVATDGTVTLAAGATKTTMPAGATTKTEVQELKDTPAVVSADAKN
ALIAGGVDATDANGAELVKMSYTDKNGKTIEGGYALKAGDKYYAADYDEATGAIKAKTTS
YTAADGTTKTAANQLGGVDGKTEVVTIDGKTYNASKAAGHDFKAQPELAEAAAKTTENPL
QKIDAALAQVDALRSDLGAVQNRFNSAITNLGNTVNNLSEARSRIEDSDYATEVSNMSRA
QILQQAGTSVLAQANQVPQNVLSLLA**MEWENIPSIQSRGLFGAIAGFIEEWENIPSIQSR**
GLFGAIAGFIE

SEQ ID NO: 344 STF2.3xH1C1

MAQVINTNSLSLLTQNNLNKSQSALGTAIERLSSGLRINSAKDDAAGQAIANRFTANIKG
LTQASRNANDGISIAQTTEGALNEINNNLQRVRELAVQSANSTNSQSDLDSIQAEITQRL
NEIDRVSGQTQFNGVKVLAQDNTLTIQVGANDGETIDIDLKQINSQTLGLDSLNVQKAYD
VKDTAVTTKAYANNGTTLDVSGLDDAAIKAATGGTNGTASVTGGAVKFDADNNKYFVTIG
GFTGADAAKNGDYEVNVATDGTVTLAAGATKTTMPAGATTKTEVQELKDTPAVVSADAKN
ALIAGGVDATDANGAELVKMSYTDKNGKTIEGGYALKAGDKYYAADYDEATGAIKAKTTS
YTAADGTTKTAANQLGGVDGKTEVVTIDGKTYNASKAAGHDFKAQPELAEAAAKTTENPL
QKIDAALAQVDALRSDLGAVQNRFNSAITNLGNTVNNLSEARSRIEDSDYATEVSNMSRA
QILQQAGTSVLAQANQVPQNVLSLLA**MEWENIPSIQSRGLFGAIAGFIEEWENIPSIQSR**
GLFGAIAGFIEEWENIPSIQSRGLFGAIAGFIE

SEQ ID NO: 345 STF2.4xH1C1

MAQVINTNSLSLLTQNNLNKSQSALGTAIERLSSGLRINSAKDDAAGQAIANRFTANIKG
LTQASRNANDGISIAQTTEGALNEINNNLQRVRELAVQSANSTNSQSDLDSIQAEITQRL
NEIDRVSGQTQFNGVKVLAQDNTLTIQVGANDGETIDIDLKQINSQTLGLDSLNVQKAYD
VKDTAVTTKAYANNGTTLDVSGLDDAAIKAATGGTNGTASVTGGAVKFDADNNKYFVTIG
GFTGADAAKNGDYEVNVATDGTVTLAAGATKTTMPAGATTKTEVQELKDTPAVVSADAKN
ALIAGGVDATDANGAELVKMSYTDKNGKTIEGGYALKAGDKYYAADYDEATGAIKAKTTS
YTAADGTTKTAANQLGGVDGKTEVVTIDGKTYNASKAAGHDFKAQPELAEAAAKTTENPL
QKIDAALAQVDALRSDLGAVQNRFNSAITNLGNTVNNLSEARSRIEDSDYATEVSNMSRA
QILQQAGTSVLAQANQVPQNVLSLLA**MEWENIPSIQSRGLFGAIAGFIEEWENIPSIQSR**
GLFGAIAGFIEEWENIPSIQSRGLFGAIAGFIEEWENIPSIQSRGLFGAIAGFIE

SEQ ID NO: 346 STF2.4xH1C2

MAQVINTNSLSLLTQNNLNKSQSALGTAIERLSSGLRINSAKDDAAGQAIANRFTANIKG
LTQASRNANDGISIAQTTEGALNEINNNLQRVRELAVQSANSTNSQSDLDSIQAEITQRL
NEIDRVSGQTQFNGVKVLAQDNTLTIQVGANDGETIDIDLKQINSQTLGLDSLNVQKAYD
VKDTAVTTKAYANNGTTLDVSGLDDAAIKAATGGTNGTASVTGGAVKFDADNNKYFVTIG
GFTGADAAKNGDYEVNVATDGTVTLAAGATKTTMPAGATTKTEVQELKDTPAVVSADAKN
ALIAGGVDATDANGAELVKMSYTDKNGKTIEGGYALKAGDKYYAADYDEATGAIKAKTTS
YTAADGTTKTAANQLGGVDGKTEVVTIDGKTYNASKAAGHDFKAQPELAEAAAKTTENPL
QKIDAALAQVDALRSDLGAVQNRFNSAITNLGNTVNNLSEARSRIEDSDYATEVSNMSRA
QILQQAGTSVLAQANQVPQNVLSLLA**GCGSEWENIPSIQSRGLFGAIAGFIEEWENIPSI**
QSRGLFGAIAGFIEEWENIPSIQSRGLFGAIAGFIEEWENIPSIQSRGLFGAIAGFIESG
C

SEQ ID NO: 347 HA maturation cleavage fragment (H5N1)

RERRRKKRGLFGAIAGFIE

SEQ ID NO: 348 STF2.4xH5C1

MAQVINTNSLSLLTQNNLNKSQSALGTAIERLSSGLRINSAKDDAAGQAIANRFTANIKG
LTQASRNANDGISIAQTTEGALNEINNNLQRVRELAVQSANSTNSQSDLDSIQAEITQRL
NEIDRVSGQTQFNGVKVLAQDNTLTIQVGANDGETIDIDLKQINSQTLGLDSLNVQKAYD
VKDTAVTTKAYANNGTTLDVSGLDDAAIKAATGGTNGTASVTGGAVKFDADNNKYFVTIG
GFTGADAAKNGDYEVNVATDGTVTLAAGATKTTMPAGATTKTEVQELKDTPAVVSADAKN
ALIAGGVDATDANGAELVKMSYTDKNGKTIEGGYALKAGDKYYAADYDEATGAIKAKTTS
YTAADGTTKTAANQLGGVDGKTEVVTIDGKTYNASKAAGHDFKAQPELAEAAAKTTENPL
QKIDAALAQVDALRSDLGAVQNRFNSAITNLGNTVNNLSEARSRIEDSDYATEVSNMSRA
QILQQAGTSVLAQANQVPQNVLSLLA**EWERERRRKKRGLFGAIAGFIEEWERERRRKKR**G
LFGAIAGFIEEWERERRRKKRGLFGAIAGFIEEWERERRRKKRGLFGAIAGFIE

SEQ ID NO: 349 HA maturation cleavage fragment (H3N2)

NVPEKQTRGIFGAIAGFIE

SEQ ID NO: 350 HA maturation cleavage fragment (H2N1/H2N2/H2N3/H2N5/H2N8/H2N9)

NVPQIESRGLFGAIAGFIE

SEQ ID NO: 351 HA maturation cleavage fragment (B strain)

PAKLLKERGFFGAIAGFLE

SEQ ID NO: 352 STF2.4xH1C1

ATGGCACAAGTAATCAACACTAACAGTCTGTCGCTGCTGACCCAGAATAACCTGAACAAA
TCCCAGTCCGCACTGGGCACCGCTATCGAGCGTCTGTCTTCTGGTCTGCGTATCAACAGC
GCGAAAGACGATGCGGCAGGTCAGGCGATTGCTAACCGTTTCACCGCGAACATCAAAGGT
CTGACTCAGGCTTCCCGTAACGCTAACGACGGTATCTCCATTGCGCAGACCACTGAAGGC
GCGCTGAACGAAATCAACAACAACCTGCAGCGTGTGCGTGAACTGGCGGTTCAGTCTGCT
AACAGCACCAACTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATCACCCAGCGCCTG
AACGAAATCGACCGTGTATCCGGCCAGACTCAGTTCAACGGCGTGAAAGTCCTGGCGCAG
GACAACACCCTGACCATCCAGGTTGGCGCCAACGACGGTGAAACTATCGATATCGATCTG
AAGCAGATCAACTCTCAGACCCTGGGTCTGGACTCACTGAACGTGCAGAAAGCGTATGAT
GTGAAAGATACAGCAGTAACAACGAAAGCTTATGCCAATAATGGTACTACACTGGATGTA
TCGGGTCTTGATGATGCAGCTATTAAAGCGGCTACGGGTGGTACGAATGGTACGGCTTCT
GTAACCGGTGGTGCGGTTAAATTTGACGCAGATAATAACAAGTACTTTGTTACTATTGGT
GGCTTTACTGGTGCTGATGCCGCCAAAAATGGCGATTATGAAGTTAACGTTGCTACTGAC
GGTACAGTAACCCTTGCGGCTGGCGCAACTAAAACCACAATGCCTGCTGGTGCGACAACT
AAAACAGAAGTACAGGAGTTAAAAGATACACCGGCAGTTGTTTCAGCAGATGCTAAAAAT
GCCTTAATTGCTGGCGGCGTTGACGCTACCGATGCTAATGGCGCTGAGTTGGTCAAAATG
TCTTATACCGATAAAAATGGTAAGACAATTGAAGGCGGTTATGCGCTTAAAGCTGGCGAT
AAGTATTACGCCGCAGATTACGATGAAGCGACAGGAGCAATTAAAGCTAAAACTACAAGT
TATACTGCTGCTGACGGCACTACCAAAACAGCGGCTAACCAACTGGGTGGCGTAGACGGT
AAAACCGAAGTCGTTACTATCGACGGTAAAACCTACAATGCCAGCAAAGCCGCTGGTCAT
GATTTCAAAGCACAACCAGAGCTGGCGGAAGCAGCCGCTAAAACCACCGAAAACCCGCTG
CAGAAAATTGATGCCGCGCTGGCGCAGGTGGATGCGCTGCGCTCTGATCTGGGTGCGGTA
CAAAACCGTTTCAACTCTGCTATCACCAACCTGGGCAATACCGTAAACAATCTGTCTGAA
GCGCGTAGCCGTATCGAAGATTCCGACTACGCGACCGAAGTTTCCAACATGTCTCGCGCG
CAGATTCTGCAGCAGGCCGGTACTTCCGTTCTGGCGCAGGCTAACCAGGTCCCGCAGAAC
GTGCTGTCTCTGTTAGCGATGGAATGGGAGAACATCCCTAGCATCCAATCTCGCGGCCTG
TTTGGCGCTATCGCGGGCTTTATCGAAGAATGGGAGAACATCCCGAGCATCCAATCTCGC
GGTCTGTTTGGTGCGATCGCTGGTTTCATCGAGGAGTGGGAGAACATTCCTAGCATTCAA
AGCCGTGGCCTGTTCGGCGCTATTGCAGGTTTTATTGAAGAATGGGAAAATATCCCGTCT
ATCCAATCCCGCGGTCTGTTCGGCGCGATCGCAGGTTTCATTGAATAATAAGCTAAGC

SEQ ID NO: 353 STF2.3xH1C1

ATGGCACAAGTAATCAACACTAACAGTCTGTCGCTGCTGACCCAGAATAACCTGAACAAA
TCCCAGTCCGCACTGGGCACCGCTATCGAGCGTCTGTCTTCTGGTCTGCGTATCAACAGC
GCGAAAGACGATGCGGCAGGTCAGGCGATTGCTAACCGTTTCACCGCGAACATCAAAGGT
CTGACTCAGGCTTCCCGTAACGCTAACGACGGTATCTCCATTGCGCAGACCACTGAAGGC
GCGCTGAACGAAATCAACAACAACCTGCAGCGTGTGCGTGAACTGGCGGTTCAGTCTGCT
AACAGCACCAACTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATCACCCAGCGCCTG
AACGAAATCGACCGTGTATCCGGCCAGACTCAGTTCAACGGCGTGAAAGTCCTGGCGCAG
GACAACACCCTGACCATCCAGGTTGGCGCCAACGACGGTGAAACTATCGATATCGATCTG
AAGCAGATCAACTCTCAGACCCTGGGTCTGGACTCACTGAACGTGCAGAAAGCGTATGAT
GTGAAAGATACAGCAGTAACAACGAAAGCTTATGCCAATAATGGTACTACACTGGATGTA
TCGGGTCTTGATGATGCAGCTATTAAAGCGGCTACGGGTGGTACGAATGGTACGGCTTCT
GTAACCGGTGGTGCGGTTAAATTTGACGCAGATAATAACAAGTACTTTGTTACTATTGGT
GGCTTTACTGGTGCTGATGCCGCCAAAAATGGCGATTATGAAGTTAACGTTGCTACTGAC
GGTACAGTAACCCTTGCGGCTGGCGCAACTAAAACCACAATGCCTGCTGGTGCGACAACT
AAAACAGAAGTACAGGAGTTAAAAGATACACCGGCAGTTGTTTCAGCAGATGCTAAAAAT
GCCTTAATTGCTGGCGGCGTTGACGCTACCGATGCTAATGGCGCTGAGTTGGTCAAAATG
TCTTATACCGATAAAAATGGTAAGACAATTGAAGGCGGTTATGCGCTTAAAGCTGGCGAT
AAGTATTACGCCGCAGATTACGATGAAGCGACAGGAGCAATTAAAGCTAAAACTACAAGT
TATACTGCTGCTGACGGCACTACCAAAACAGCGGCTAACCAACTGGGTGGCGTAGACGGT
AAAACCGAAGTCGTTACTATCGACGGTAAAACCTACAATGCCAGCAAAGCCGCTGGTCAT
GATTTCAAAGCACAACCAGAGCTGGCGGAAGCAGCCGCTAAAACCACCGAAAACCCGCTG
CAGAAAATTGATGCCGCGCTGGCGCAGGTGGATGCGCTGCGCTCTGATCTGGGTGCGGTA
CAAAACCGTTTCAACTCTGCTATCACCAACCTGGGCAATACCGTAAACAATCTGTCTGAA
GCGCGTAGCCGTATCGAAGATTCCGACTACGCGACCGAAGTTTCCAACATGTCTCGCGCG
CAGATTCTGCAGCAGGCCGGTACTTCCGTTCTGGCGCAGGCTAACCAGGTCCCGCAGAAC
GTGCTGAGCCTGTTAGCGATGGAATGGGAAAATATCCCTAGCATCCAATCTCGCGGTCTG
TTCGGTGCTATTGCTGGCTTCATCGAGGAATGGGAGAACATCCCATCTATTCAGTCTCGC
GGCCTGTTTGGTGCGATCGCGGGTTTTATTGAGGAATGGGAAAACATTCCAAGCATTCAG
TCACGTGGTCTTTTCGGCGCCATCGCTGGTTTTATCGAATGATAAGCTTAGCCCAAGG

SEQ ID NO: 354 STF2.2xH1C1

ATGGCACAAGTAATCAACACTAACAGTCTGTCGCTGCTGACCCAGAATAACCTGAACAAA
TCCCAGTCCGCACTGGGCACCGCTATCGAGCGTCTGTCTTCTGGTCTGCGTATCAACAGC
GCGAAAGACGATGCGGCAGGTCAGGCGATTGCTAACCGTTTCACCGCGAACATCAAAGGT
CTGACTCAGGCTTCCCGTAACGCTAACGACGGTATCTCCATTGCGCAGACCACTGAAGGC
GCGCTGAACGAAATCAACAACAACCTGCAGCGTGTGCGTGAACTGGCGGTTCAGTCTGCT
AACAGCACCAACTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATCACCCAGCGCCTG
AACGAAATCGACCGTGTATCCGGCCAGACTCAGTTCAACGGCGTGAAAGTCCTGGCGCAG
GACAACACCCTGACCATCCAGGTTGGCGCCAACGACGGTGAAACTATCGATATCGATCTG
AAGCAGATCAACTCTCAGACCCTGGGTCTGGACTCACTGAACGTGCAGAAAGCGTATGAT
GTGAAAGATACAGCAGTAACAACGAAAGCTTATGCCAATAATGGTACTACACTGGATGTA
TCGGGTCTTGATGATGCAGCTATTAAAGCGGCTACGGGTGGTACGAATGGTACGGCTTCT
GTAACCGGTGGTGCGGTTAAATTTGACGCAGATAATAACAAGTACTTTGTTACTATTGGT
GGCTTTACTGGTGCTGATGCCGCCAAAAATGGCGATTATGAAGTTAACGTTGCTACTGAC
GGTACAGTAACCCTTGCGGCTGGCGCAACTAAAACCACAATGCCTGCTGGTGCGACAACT
AAAACAGAAGTACAGGAGTTAAAAGATACACCGGCAGTTGTTTCAGCAGATGCTAAAAAT
GCCTTAATTGCTGGCGGCGTTGACGCTACCGATGCTAATGGCGCTGAGTTGGTCAAAATG
TCTTATACCGATAAAAATGGTAAGACAATTGAAGGCGGTTATGCGCTTAAAGCTGGCGAT
AAGTATTACGCCGCAGATTACGATGAAGCGACAGGAGCAATTAAAGCTAAAACTACAAGT
TATACTGCTGCTGACGGCACTACCAAAACAGCGGCTAACCAACTGGGTGGCGTAGACGGT
AAAACCGAAGTCGTTACTATCGACGGTAAAACCTACAATGCCAGCAAAGCCGCTGGTCAT

GATTTCAAAGCACAACCAGAGCTGGCGGAAGCAGCCGCTAAAACCACCGAAAACCCGCTG
CAGAAAATTGATGCCGCGCTGGCGCAGGTGGATGCGCTGCGCTCTGATCTGGGTGCGGTA
CAAAACCGTTTCAACTCTGCTATCACCAACCTGGGCAATACCGTAAACAATCTGTCTGAA
GCGCGTAGCCGTATCGAAGATTCCGACTACGCGACCGAAGTTTCCAACATGTCTCGCGCG
CAGATTCTGCAGCAGGCCGGTACTTCCGTTCTGGCGCAGGCTAACCAGGTCCCGCAGAAC
GTGCTGAGCCTGTTAGCGATGGAATGGGAAAATATCCCTAGCATCCAATCTCGCGGTCTG
TTCGGTGCTATTGCTGGCTTCATCGAGGAATGGGAGAACATCCCATCTATTCAGTCTCGC
GGCCTGTTTGGTGCGATCGCGGGTTTTATTGAGTGATAAGCTTAGCCCAAGG

SEQ ID NO: 355 STF2.1xH1C1

ATGGCACAAGTAATCAACACTAACAGTCTGTCGCTGCTGACCCAGAATAACCTGAACAAA
TCCCAGTCCGCACTGGGCACCGCTATCGAGCGTCTGTCTTCTGGTCTGCGTATCAACAGC
GCGAAAGACGATGCGGCAGGTCAGGCGATTGCTAACCGTTTCACCGCGAACATCAAAGGT
CTGACTCAGGCTTCCCGTAACGCTAACGACGGTATCTCCATTGCGCAGACCACTGAAGGC
GCGCTGAACGAAATCAACAACAACCTGCAGCGTGTGCGTGAACTGGCGGTTCAGTCTGCT
AACAGCACCAACTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATCACCCAGCGCCTG
AACGAAATCGACCGTGTATCCGGCCAGACTCAGTTCAACGGCGTGAAAGTCCTGGCGCAG
GACAACACCCTGACCATCCAGGTTGGCGCCAACGACGGTGAAACTATCGATATCGATCTG
AAGCAGATCAACTCTCAGACCCTGGGTCTGGACTCACTGAACGTGCAGAAAGCGTATGAT
GTGAAAGATACAGCAGTAACAACGAAAGCTTATGCCAATAATGGTACTACACTGGATGTA
TCGGGTCTTGATGATGCAGCTATTAAAGCGGCTACGGGTGGTACGAATGGTACGGCTTCT
GTAACCGGTGGTGCGGTTAAATTTGACGCAGATAATAACAAGTACTTTGTTACTATTGGT
GGCTTTACTGGTGCTGATGCCGCCAAAAATGGCGATTATGAAGTTAACGTTGCTACTGAC
GGTACAGTAACCCTTGCGGCTGGCGCAACTAAAACCACAATGCCTGCTGGTGCGACAACT
AAAACAGAAGTACAGGAGTTAAAAGATACACCGGCAGTTGTTTCAGCAGATGCTAAAAAT
GCCTTAATTGCTGGCGGCGTTGACGCTACCGATGCTAATGGCGCTGAGTTGGTCAAAATG
TCTTATACCGATAAAAATGGTAAGACAATTGAAGGCGGTTATGCGCTTAAAGCTGGCGAT
AAGTATTACGCCGCAGATTACGATGAAGCGACAGGAGCAATTAAAGCTAAAACTACAAGT
TATACTGCTGCTGACGGCACTACCAAAACAGCGGCTAACCAACTGGGTGGCGTAGACGGT
AAAACCGAAGTCGTTACTATCGACGGTAAAACCTACAATGCCAGCAAAGCCGCTGGTCAT
GATTTCAAAGCACAACCAGAGCTGGCGGAAGCAGCCGCTAAAACCACCGAAAACCCGCTG
CAGAAAATTGATGCCGCGCTGGCGCAGGTGGATGCGCTGCGCTCTGATCTGGGTGCGGTA
CAAAACCGTTTCAACTCTGCTATCACCAACCTGGGCAATACCGTAAACAATCTGTCTGAA
GCGCGTAGCCGTATCGAAGATTCCGACTACGCGACCGAAGTTTCCAACATGTCTCGCGCG
CAGATTCTGCAGCAGGCCGGTACTTCCGTTCTGGCGCAGGCTAACCAGGTCCCGCAGAAC
GTGCTGAGCCTGTTAGCGATGGAATGGGAAAATATCCCTAGCATCCAATCTCGCGGTCTG
TTCGGTGCTATTGCTGGCTTCATCGAGTGATAAGCTTAGCCCAAGG

SEQ ID NO: 356 STF2.4xH1C1

ATGGCACAAGTAATCAACACTAACAGTCTGTCGCTGCTGACCCAGAATAACCTGAACAAA
TCCCAGTCCGCACTGGGCACCGCTATCGAGCGTCTGTCTTCTGGTCTGCGTATCAACAGC
GCGAAAGACGATGCGGCAGGTCAGGCGATTGCTAACCGTTTCACCGCGAACATCAAAGGT
CTGACTCAGGCTTCCCGTAACGCTAACGACGGTATCTCCATTGCGCAGACCACTGAAGGC
GCGCTGAACGAAATCAACAACAACCTGCAGCGTGTGCGTGAACTGGCGGTTCAGTCTGCT
AACAGCACCAACTCCCAGTCTGACCTCGACTCCATCCAGGCTGAAATCACCCAGCGCCTG
AACGAAATCGACCGTGTATCCGGCCAGACTCAGTTCAACGGCGTGAAAGTCCTGGCGCAG
GACAACACCCTGACCATCCAGGTTGGCGCCAACGACGGTGAAACTATCGATATCGATCTG
AAGCAGATCAACTCTCAGACCCTGGGTCTGGACTCACTGAACGTGCAGAAAGCGTATGAT
GTGAAAGATACAGCAGTAACAACGAAAGCTTATGCCAATAATGGTACTACACTGGATGTA
TCGGGTCTTGATGATGCAGCTATTAAAGCGGCTACGGGTGGTACGAATGGTACGGCTTCT
GTAACCGGTGGTGCGGTTAAATTTGACGCAGATAATAACAAGTACTTTGTTACTATTGGT
GGCTTTACTGGTGCTGATGCCGCCAAAAATGGCGATTATGAAGTTAACGTTGCTACTGAC
GGTACAGTAACCCTTGCGGCTGGCGCAACTAAAACCACAATGCCTGCTGGTGCGACAACT

AAAACAGAAGTACAGGAGTTAAAAGATACACCGGCAGTTGTTTCAGCAGATGCTAAAAAT
GCCTTAATTGCTGGCGGCGTTGACGCTACCGATGCTAATGGCGCTGAGTTGGTCAAAATG
TCTTATACCGATAAAAATGGTAAGACAATTGAAGGCGGTTATGCGCTTAAAGCTGGCGAT
AAGTATTACGCCGCAGATTACGATGAAGCGACAGGAGCAATTAAAGCTAAAACTACAAGT
TATACTGCTGCTGACGGCACTACCAAAACAGCGGCTAACCAACTGGGTGGCGTAGACGGT
AAAACCGAAGTCGTTACTATCGACGGTAAAACCTACAATGCCAGCAAAGCCGCTGGTCAT
GATTTCAAAGCACAACCAGAGCTGGCGGAAGCAGCCGCTAAAACCACCGAAAACCCGCTG
CAGAAAATTGATGCCGCGCTGGCGCAGGTGGATGCGCTGCGCTCTGATCTGGGTGCGGTA
CAAAACCGTTTCAACTCTGCTATCACCAACCTGGGCAATACCGTAAACAATCTGTCTGAA
GCGCGTAGCCGTATCGAAGATTCCGACTACGCGACCGAAGTTTCCAACATGTCTCGCGCG
CAGATTCTGCAGCAGGCCGGTACTTCCGTTCTGGCGCAGGCTAACCAGGTCCCGCAGAAC
GTGCTGTCTCTGTTAGCGGGTTGTGGTTCCGAGTGGGAAAATATTCCGTCTATCCAGAGC
CGTGGTCTGTTCGGCGCAATTGCTGGCTTCATTGAAGAATGGGAAAACATCCCGTCCATC
CAGAGCCGTGGCCTGTTCGGCGCCATTGCTGGTTTCATCGAGGAATGGGAAAACATTCCG
TCCATCCAGTCCCGCGGTCTGTTTGGCGCTATCGCCGGTTTCATTGAGGAATGGGAAAAT
ATCCCTTCCATCCAGTCTCGTGGTCTGTTCGGCGCGATTGCAGGCTTTATCGAATCTGGT
TGCTAATAAGCTAAGC

SEQ ID NO: 357 Bacterial lipoprotein of *E. coli*
MKATKLVLGAVILGSTLLAGCSSNAKIDQLSSDVQTLNAKVDQLSNDVNAMRSDVQAAKDDAARA
NQRLDNMATKYRK

SEQ ID NO: 358 H1C1 native peptide
SLWSEENIPSIQSRGLFGAIAGFIEE

EXAMPLE 15: FLAGELLIN-M2e FUSION PROTEINS

M2e is conserved across multiple influenza A subtypes (also referred to herein as "strain"). M2e is at least a portion of the M2 protein, in particular, a 24 amino-terminus (also referred to herein as an "ectodomain") of the M2 protein. The M2 ectodomain is relatively small amino acid sequence (24 amino acids) compared to HA (about 566 amino acids) and NA (about 469 amino acids). The M2e sequence of exemplary avian influenza A isolates differs from that of human isolates, but is highly-conserved among the avian isolates (see, for example, SEQ ID NOS: 544-556, 570 and 573-578). Four tandem copies of M2e fused to the carboxy terminus of a flagellin STF2 (full-length or STF2 hinge region-deleted) were generated. The STF2 without the hinge region is also referred to herein as "STF2Δ."

Construction of Fusion Protein

The carboxy-terminal fusion of the synthetic 4xM2e sequence (4 consecutive 24 amino acid sequences) with STF2 was constructed as follows. The pET24A vector was purchased from Novagen, San Diego, CA. The strategy employed the

Seamless Cloning Kit (Catalog number 214400) from Stratagene (La Jolla, CA www.stratagene.com) performed by DNA 2.0 Inc. (Menlo Park, CA). The gene encoding the fusion protein was in pDrive 4xM2e G00448 and was used as a PCR template for insert preparation for construction of the C-terminal fusion expression construct with STF2. The synthetic 4xM2e construct pDrive 4xM2e G00448 was used as a template for PCR as outlined in the Seamless Cloning Kit (Catalog number 214400) from Stratagene (La Jolla, CA). The expected product from this amplification includes the 318 bp and the restriction enzyme sites incorporated into the oligonucleotides used to amplify this insert. The procedure was as follows:

PCR conditions

1 μL -20 ng of pDrive 4xM2e G00448

5 μL of 10x cloned Pfu polymerase buffer

1 μL of 40 mM dNTP mix

1 μL -10 pmol of forward primer 4xM2eforbs1

1 μL -10 pmol of reverse primer 4xM2erevwsto

40 μL ddH_2O

Immediately before starting the thermal cycling 1 μL of *PfuTurbo* DNA Polymerase the following were added.

4xM2eforbs1 primer sequence:

5'–CGCTCTTCAMTGAGCTTGCTGACTGAGGTTGAGACCCCGATTC (SEQ ID NO: 566)

4xM2erevwsto primer sequence:

5'-

CGCTCTTCACGCTTATTATCTAGACGGGTCTGAGCTATCGTTAGAGCGAG (SEQ ID NO: 567)

This reaction was cycled as follows on a Thermo Hybaid PxE thermal cycler (Waltham, MA).

Initial cycle

Temperature	Duration
95°	3 minutes
65°	1 minute
72°	1 minute

Subsequent nine cycles

Temperature	Duration
95°	45 seconds
65°	35 seconds
72°	1 minute

At this point the following was added to each reaction.

5 μL of 10x cloned Pfu polymerase buffer

1 μL of 5-methyl dNTP mix

44 μL ddH_2O

Subsequently the following thermal cycling was repeated five times.

Temperature	Duration
95°	45 seconds
65°	35 seconds
72°	1 minute

The 100 μL product was brought to a volume of 300 μL by the addition of TE buffer. The resulting product was phenol chloroform (Invitrogen Carlsbad, CA-Catalog number 15593-031) extracted once and chloroform extracted once. The amplification product was then ethanol precipitated by addition of 30 μL of Sodium

acetate buffer pH 5.2 and 750 µL of 100% Ethanol. The DNA pellet was washed twice with 300 µL 70% Ethanol allowed to air dry for ten minutes and then resuspended in 50 µL TE buffer.

Amplification of Vector STF2 in pET24.

The previously constructed pET24a/STF2.M2e construct was used as a template for PCR as outlined in the Seamless Cloning Kit (Catalog number 214400) from Stratagene (La Jolla, CA). The expected product from this amplification includes the whole of the pET24 plasmid plus the STF2 sequences but does not include the single copy of M2E that exists in this construct. The procedure was as follow:

PCR conditions

1 µL -40 ng of STF2.M2E pET22-2

5 µL of 10x cloned Pfu polymerase buffer

1 µL of 40 mM dNTP mix

1 µL -10 pmol of primer 4xMECpET24

1 µL -10 pmol of primer 4xM2eC-STF2

40 µL ddH_2O

Immediately before starting the thermal cycling the following were added:

1 µL of *PfuTurbo* DNA Polymerase

4xMECpET24 primer sequence:

5'-GCTCTTCAGCGGCTGAGCAATAACTAGCATAACCCCTTGGG (SEQ ID NO: 568)

4xM2eC-STF2 primer sequence:

5'-CGCTCTTCACAGACGTAACAGAGACAGCACGTTCTGCGG (SEQ ID NO: 569)

This reaction was cycled as follows on a Thermo Hybaid PxE thermal cycler (Waltham, MA).

Initial cycle

Temperature	Duration
95°	3 minutes
65°	1 minute
72°	18 minutes

Subsequent nine cycles

Temperature	Duration
95°	45 seconds
65°	35 seconds
72°	18 minutes

At this point the following was added to each reaction.

5 μL of 10x cloned Pfu polymerase buffer
1 μL of 5-methyl dNTP mix
44 μL ddH_2O

Subsequently the following thermal cycling was repeated five times.

Temperature	Duration
95°	45 seconds
65°	35 seconds
72°	18 minutes

The 100 μL product was brought to a volume of 300 μL by the addition of TE buffer. The resulting product was phenol chloroform (Invitrogen Carlsbad, CA-Catalog number 15593-031) extracted once and chloroform extracted once. The amplification product was then ethanol precipitated by addition of 30 μL of Sodium acetate buffer pH 5.2 and 750 μL of 100% Ethanol. The DNA pellet was washed twice with 300 μL 70% Ethanol allowed to air dry for ten minutes and then resuspended in 50 μL TE buffer.

Digestion and ligation of Vector and Insert amplifications

Eam 1104 I digests were set up separately for vector and insert as follows:

30 µL of amplified product after ethanol precipitation

5 µL of 10x Universal buffer (Supplied with Seamless Cloning Kit)

4 µL *Eam* 1104 I restriction enzyme (Supplied with Seamless Cloning Kit)

11 µL ddH_2O

Digests were mixed gently and incubated at 37°C for one hour and ligation reactions of vector and insert products were prepared as above performed as follows (Reagents supplied with Seamless Cloning Kit):

Ingredients added in order listed:

9 µL ddH_2O

5 µL of *Eam* 1104 I digested 4xM2e amplified insert

5 µL of *Eam* 1104 I digested STF2.M2E pET22-2 amplified vector

2 µL 10x Ligase buffer

2 µL 10 mM rATP

1 µL T4 DNA Ligase (diluted from stock 1:16)

1 µL *Eam* 1104 I restriction enzyme

The ligation reactions were mixed gently and incubated for 30 minutes at 37°C. The ligations were then stored on ice until transformed into XL-10 competent cells (Stratagene Catalog number 200314) later than same day.

Transformation of Ligation into XL-10 Competent Cells

Eppendorf tubes were chilled for ten minutes while the XL-10 (Stratagene Catalog number 200314) competent cells thawed on ice.

50 µL of competent cells were aliquoted from the stock tube per ligation.

2 µL of β-mercaptoethanol stock which is provided with the XL-10 cells.

This mixture was incubated for ten minutes on ice gently mixing every 2 minutes. Seamless cloning ligation reaction (4µl) was added, swirled gently and then incubated on ice for 30 minutes. The tubes were heat shocked for 35 seconds at 42°C in a water bath. The tubes were incubated on ice for at least two minutes. SOC medium (400 µL) were added to the cells and incubated for one hour at 37°C with agitation.

Two LB agar kanamycin (50μg/mL) plates are used to plate 200 μL and 10 μL of the transformed cells and allowed to grow overnight.

Screening of Kanamycin Resistant Clones

Recombinant candidates were grown up for minipreps in Luria Broth containing Kanamycin (25 ug/mL) and extracted using the QIAprep Spin Miniprep Kit (Qiagen Valencia, CA Catalog Number 27106). Candidate clones were screened by restriction enzymes (New England Biolabs Beverly, MA) and positive clones were grown up in 100 mL of Luria Broth containing kanamycin (25 ug/mL) and extracted using the Qiagen HiSpeed Plasmid Midi Kit (Catalog number 12643). These clones were submitted to GENEWIZ (North Brunswick, NJ) for sequencing.

Production and Purification of STF2.4xM2e Fusion Protein

STF2.4xM2e in *E. coli* BLR(DE3)pLysS host (Novagen, San Diego, CA, Catalog #69053) was retrieved from glycerol stock and scaled up to 5 L. Cells were grown in LB medium containing 15 μg/ml Kanamycin and 12.5 μg/ml Teteracycline to OD_{600} = 0.4 and induced with 1 mM IPTG for 3 h at 37°C. The cells were harvested by centrifugation (7000 rpm x 7 minutes in a Sorvall RC5C centrifuge) and resuspended in 2x PBS, 1% glycerol, DNAse, 1 mM PMSF, protease inhibitor cocktail and 1 mg/ml lysozyme. The suspension was passed through a microfluidizer to lyse the cells. The lysate was centrifuged (45,000 g for one hour in a Beckman Optima L ultracentrifuge) to separate the soluble fraction from inclusion bodies. Protein was detected by SDS-PAGE in the soluble and insoluble fractions.

The soluble fraction was applied to Sepharose Q resin in the presence of high salt via batch method to reduce DNA, endotoxin, and other contaminants. The flow through containing the protein of interest was loaded onto 30 ml Q Sepharose column (Amersham Biosciences). Bound protein was eluted using a linear gradient from Buffer A to B. (Buffer A: 100 mM Tris-Cl, pH 8.0. Buffer B: 100 mM Tris-Cl, 1 M NaCl, pH 8.0). Eluted protein was further purified using a 45 ml Source Q column that provided greater resolution needed to resolve contaminating proteins. Bound protein was eluted with a linear gradient from Buffer A to B (Buffer A: 100 mM Tris-Cl, pH 8.0 Buffer B: 100 mM Tris-Cl, 1 M NaCl, pH 8.0).

Final purification of protein was completed using Superdex-200 gel filtration chromatography. The column was developed with 100 mM Tris, 150 mM NaCl and 1% glycerol plus 1% Na-deoxycholate to remove the LPS. Buffer exchange was carried out using overnight dialysis against buffer containing 50 mM Tris, 100 mM NaCl and 1% glycerol was done to remove Na-deoxycholate. Protein concentration was determined by the MicroBCA Protein Assay Reagent Kit (Pierce Biotechnology). Purified preparations of STF2.4xM2e yielded a single band visible with Coomassie stain that migrated with an apparent molecular weight of about 64 kDa on 12% SDS polyacrylamide gels.

EXAMPLE 16: EXPRESSION AND PURIFICATION OF FLAGELLIN (STF2 AND STF2Δ) FUSION PROTEIN CONSTRUCTS ENCODING INFLUENZA A M2 ECTODOMAIN SEQUENCES

The consensus M2e sequences from several influenza A strains of human and avian origin are depicted in SEQ ID NOS: 544-556, 570 and 573-578. To facilitate the cloning of the M2e sequence, two vector cassettes, pMT/STF2 and pMT/STF2Δ, each containing a multiple cloning site (MCS) were generated (See Figures 17A and 17B). To generate pMT/STF2, the 1.5 kb gene encoding full length flagellin of *Salmonella typhimurium* fljb type 2, or STF2, was fused to the Ig binding protein (BIP) secretion signal of pMTBIP/V5-His vector (Invitrogen Corporation, Carlsbad, CA) for expression in *Drosophila*. The BiP sequence is included at the 5' end of the construct as a secretion signal for expression in *Drosophila*. A chemically-synthesized 4xM2e gene representing the H1, H2 and H3 consensus sequence, SEQ ID NO: 544, was cloned into the MCS of pMT/STF2 to create *pMT/STF2.4xM2e(H1)*.

A similar strategy prophetically is employed to clone two H5-associated M2e sequences, SLLTEVETPTRNEWECRCSDSSDP (SEQ ID NO: 553) (A/Viet Nam/1203/2004) and SLLTEVETLTRNGWGCRCSDSSDP (SEQ ID NO: 552) (A/Hong Kong/156/97). Codon-optimized chemically synthesized genes containing four tandemly repeated copies of the indicated H5-associated M2e sequence prophetically are cloned into pMT/STF2 to generate *STF2.4xM2e(H5VN)* and *STF2.4xM2e(H5HK)*, respectively. To generate a construct that contains multiple

M2e forms, the heterologous 4xM2e sequence(s) prophetically are inserted into either of the primary constructs.

"Heterologous sequences," as used herein, means sequences from different species. For example, the H1 sequence is a human sequence and the H5 sequence is an avian sequence. Thus, the H1 and H5 sequences are heterologous sequences (e.g., SLLTEVETPTRNEWESRSSDSSDPLESLLTEVETPTRNEWESRSSDSSDPESSL LTEVETPTRNEWESRSSDSSDPGSSLLTEVETPTRNEWESRSSDSSDP (SEQ ID NO: 597), encoded by tctctgctgactgaagtagaaactccaacgcgtaatgaatgggaatcccgttctagcgactcctctgatcctctcgagtccc tgctgacggaggttgaaaccccgacccgcaacgagtgggaaagccgttcctccgattcctctgatccggagagcagcc tgctgaccgaggtagaaaccccgacccgtaatgagtgggaatctcgctcctctgattcttctgaccgggatcctctctgc tgaccgaagtggagactccgactcgcaacgaatgggagagccgttcttctgactcctctgacccg (SEQ ID NO: 598).

Primary constructs comprise at least one pathogen-associated molecular pattern (e.g., STF2, STF2Δ) and at least a portion of at least one integral membrane protein (e.g., M2e, such as SEQ ID NOS: 510 and 544). If there is more than one integral membrane in a primary construct, the integral membrane proteins are from the same species.

A heterologous construct includes at least two integral membrane proteins such as H1 (human) and H5 (avian), for example, in SEQ ID NOS: 583 and 584.

To generate pMT/STF2Δ, the hyper-variable region that spans amino acids 170 to 415 of the full-length flagellin gene of SEQ ID NO: 499 was deleted and replaced with a short (10 amino acid) flexible linker (GAPVDPASPW, SEQ ID NO: 594) designed to facilitate interactions of the amino and carboxy terminal sequences necessary for TLR5 signaling. The protein expressed from this construct retains potent TLR5 activity whether expressed alone or in fusion with test antigen. Thus, a second series of M2e constructs prophetically is generated based on pMT/STF2Δ. *Drosophila* Dmel-2 cells (Invitrogen Corporation, Carlsbad, CA) grown at room temperature in Schneider's medium supplemented with 10% FBS and antibiotics prophetically is transfected with the constructs described above using Cellfectin reagent (Invitrogen) according to the manufacturer's instructions. Twenty-four

hours post transfection, cells prophetically is induced with 0.5 mM $CuSO_4$ in medium lacking FBS and incubated for an additional 48 hours. Conditioned media (CM) prophetically is harvested from induced cultures and screened for protein expression by SDS-PAGE and Western blot analyses using anti-flagellin and anti-M2e specific antibodies. The identity, TLR bioactivity of the fusion protein, antigenicity assessed by ELISA and *in vivo* mouse studies for immunogenicity prophetically is performed.

EXAMPLE 17: CONSTRUCTION AND EXPRESSION OF FLAGELLIN-HEMAGLUTININ (HA) CONSTRUCTS

The gene encoding HA from genomic DNA from the in-house laboratory strain PR8, an attenuated derivative of A/Puerto Rico/8/34 was isolated (SEQ ID NO: 565, encoding SEQ ID NO: 564). The gene was fused to the STF2Δ cassette that has been previously constructed in pPICZΔ generating STF2Δ.HAPR8 (SEQ ID NO: 560, encoding SEQ ID NO: 559) (See Figure 42). Purified recombinant protein was tested for immunogenicity and efficacy in BALB/c mice. The gene encoding H5N1 of the A/Vietnam/1203/04 strain was custom synthesized and fused to STF2Δ cassette generating STF2Δ.HAH5 (SEQ ID NO: 558, encoding SEQ ID NO: 557). Both human and avian HA constructs were transformed into *Pichia pastoris* strains GS115 and X-33 (Invitrogen Corporation, Carlsbad, CA). Selected clones were screened for expression by fractionation on SDS-PAGE gel and staining by Coommassie Blue and Western blot analysis using anti-HA and anti-flagellin antibodies.

EXAMPLE 18: GENERATION OF A PAM3CYS FUSION PROTEIN

M2e (SEQ ID NO: 544) was chemically coupled to a tri-palmitoylcysteine (Pam3Cys) moiety through the amino terminal serine residue of the peptide. The structure of the fusion protein (Pam3Cys.M2e) is shown in Figure 39. The chemical name for Pam3Cys.M2e is [Palmitoyl-Cys((RS)-2,3-di(palmitoyloxy)-propyl)-Ser-Leu-Leu-Thr-Glu-Val-Glu-Thr-Pro-Ile-Arg-Asn-Glu-Trp-Gly-Ser-Arg-Ser-Asn-Asp-Ser-Ser-Asp-Pro-OH acetate salt]. The molecular mass of Pam3Cys.M2e is 3582.3 daltons.

Pam3Cys.M2e was synthesized using a solid phase peptide synthesis methodology based on a well established Fmoc-strategy (Houben-Weyl, 2004. Synthesis of peptides and peptidomimetics, Vol. 22, Georg Thieme Verlag Stuttgart, NY). The synthetic scheme and manufacturing process for Pam3Cys.M2e is diagrammed in the flow chart below. The Pam3Cys.M2e is a fusion protein (chemically linked) and is also referred to herein as a "lipidated peptide."

The first step in the synthesis included solid phase peptide synthesis. The amino acid sequence of Pam3Cys.M2e was assembled on an H-Pro-2-chlorotrityl chloride resin by solid phase peptide synthesis. This resin is highly suitable for the formation of peptides with the Fmoc-strategy. The peptide chain was elongated by successive coupling of the amino acid derivatives. Each coupling step was preceded by an Fmoc-deprotection step and both steps were accompanied by repeated washing of the resin. After coupling of the last amino acid derivative, the final Fmoc-deprotection step was performed. Finally, the peptide resin was washed and dried under reduced pressure. During solid phase peptide synthesis color indicator tests were performed for each step to monitor the completion of the Fmoc-cleavage and the subsequent coupling of the amino acid derivatives.

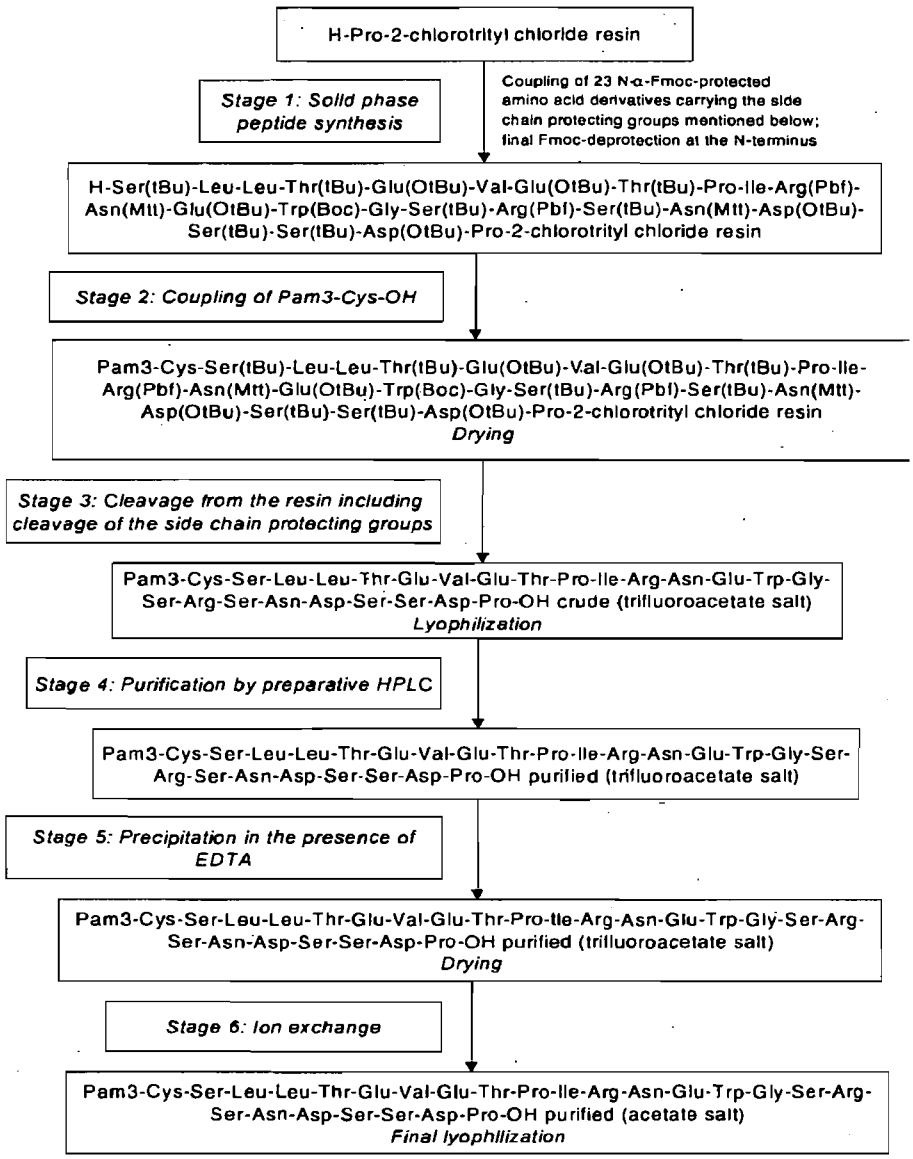
Stage 2 of the synthesis included coupling of Pam3Cys-OH. Pam3Cys-OH was pre-activated with N,N'-dicyclohexyl-carbodiimide (DCCI) in the presence of 1-hydroxybenzotriazole (HOBt). The resulting solution was filtered and added to the peptide resin. At the end of the reaction time the peptide resin was washed and dried under reduced pressure. Color indicator tests were performed to control the coupling of Pam3Cys-OH.

Stage 3 of the synthesis included cleavage from the resin including cleavage of the side chain protecting groups. The peptide resin was treated with trifluoroacetic acid (TFA). The product was precipitated from the reaction mixture and lyophilized.

Stage 4 of the synthesis included purification by preparative reverse phase HPLC. The crude material obtained from Stage 3 was purified by preparative HPLC on a reverse phase column using a TFA system. The fractions were collected, checked by analytical HPLC and pooled accordingly. Pooled fractions from the TFA runs were lyophilized.

Stage 5 of the synthesis included precipitation in the presence of EDTA. The purified material from Stage 4 was precipitated from an aqueous solution of EDTA. The product was filtered off and dried under reduced pressure.

Stage 6 of the synthesis included ion exchange chromatography. The last stage of manufacturing Pam3Cys.M2e was the exchange from the trifluoroacetate salt into the acetate salt by ion exchange. The material from Stage 5 was loaded onto an ion exchange column and eluted with acetic acid. Fractions were checked by thin layer chromatography and the combined product-containing fractions were filtered and lyophilized to yield the final product.



Pam3-Cys-OH = Palmitoyl-Cys((RS)-2,3-di(palmitoyloxy)-propyl)-OH

The purity specification for the Pam3Cys.M2e drug substance was ≥ 80% by RP-HPLC. The specification was based on the purity achieved with three non-GMP lots of Pam3Cys.M2e made from the same GMP batch of M2e-peptide intermediate resin. The purity of the three non-GMP lots of Pam3Cys.M2e was 80.2%, 80.3% and 80.8%, for lots D.001.Pam3Cys.M2e, D.002.Pam3Cys.M2e and D.003.Pam3Cys.M2e, respectively.

EXAMPLE 19: IMMUNOGENICITY

MATERIALS and METHODS

SYNTHESIS AND PURIFICATION OF PAM3CYS.M2E

Pam3Cys.M2e was prepared by Genemed Synthesis and Bachem using solid phase synthesis methodologies and FMOC chemistry as described above. Mass spectroscopy analysis was used to verify the molecular weight of the final product.

ENDOTOXIN ASSAY

Endotoxin levels of the STF2.4xM2e and the Pam3Cys.M2e were measured using the QCL-1000 Quantitative Chromogenic LAL test kit (BioWhittaker #50-648U), following the manufacturer's instructions for the microplate method.

TLR5 BIOACTIVITY ASSAY

HEK293 cells constitutively express TLR5 and secrete several soluble factors, including IL-8, in response to TLR5 signaling. HEK293 cells were seeded in 96-well microplates (50,000 cells/well) and test proteins were added and incubated overnight. The next day, the conditioned medium was harvested, transferred to a clean 96-well microplate and frozen at -20°C. After thawing, the conditioned medium was assayed for the presence of IL-8 in a sandwich ELISA using an anti-human IL-8 matched antibody pair (Pierce, #M801E and #M802B) following the manufacturer's instructions. Optical density was measured using a microplate spectrophotometer (FARCyte, Amersham). Results are reported as pg of IL8 per ml as determined by inclusion of a standard curve for IL8 in the assay.

TLR2 BIOACTIVITY ASSAY

RAW264.7 cells (ATCC) express TLR2 and secrete several soluble factors, including TNFα, in response to TLR2 signaling. RAW264.7 cells were seeded in 96-well microplates (50,000 cells/well), test compounds were added and incubated overnight. The next day, the conditioned medium was harvested, transferred to a clean 96-well microplate and frozen at -20°C. After thawing, the conditioned medium was assayed for the presence of TNFα in a sandwich ELISA using an anti-mouse TNFα matched antibody pair (Pierce) following the manufacturer's

instructions. Optical density was measured using a microplate spectrophotometer (FARCyte, Amersham). Results are reported as ng of TNF per ml as determined by reference to a standard curve for TNF included in the assay.

MOUSE IMMUNOGENICITY

Female BALB/c mice (National Cancer Institute) were used at the age of about 6-8 weeks. Mice were divided into groups of 5 to 10 mice per group, and immunized subcutaneously on each side of the base of the tail on days 0 and 21 with the indicated concentrations of STF2.4xM2e or Pam3Cys.M2e fusion protein. On days 10 (primary) and 28 (boost), individual mice were bled by retro-orbital puncture. Sera were harvested by clotting and centrifugation of the heparin-free blood samples.

MOUSE SERUM ANTIBODY DETERMINATION

M2e-specific IgG levels were determined by ELISA. 96-well ELISA plates were coated overnight at 4°C with 100 µl /well of a 5 µg/ml solution of the M2e peptide in PBS. Plates were blocked with 200 µl/well of Assay Diluent Buffer (ADB; BD Pharmingen) for one hour at room temperature. The plates were washed three times in PBS containing 0.05% Tween-20 (PBS-T). Dilutions of the sera in ADB were added (100 µl/well) and the plates were incubated overnight at 4°C. The plates were washed three times with PBS-T. Horse radish peroxidase, or HRP-labeled goat anti-mouse IgG antibodies (Jackson Immunochemical) diluted in ADB were added (100 µl/well) and the plates were incubated at room temperature for 1 hour. The plates were washed three times with PBS-T. After adding TMB Ultra substrate (3,3',5,5'-tetramentylbenzidine; Pierce) and monitoring color development, the O.D. 450 was measured on a Tecan Farcyte microspectrophotometer.

RABBIT IMMUNOGENICITY

Female and male NZW rabbits (Covance Research Products) were used at the age of about 13-17 weeks. Rabbits were divided into groups of 3 male and 3 female per group, and immunized *i.m.* on alternating thighs on days 0 and 21 and 42 with the indicated concentrations of Pam3Cys.M2e peptide or STF2.4xM2e fusion

protein. Animals were bled on day -1 (prebleed), 14 (primary) and 28 and 42 (boost). Sera were prepared by clotting and centrifugation of samples.

RABBIT SERUM ANTIBODY DETERMINATION

M2e-specific IgG levels were determined by ELISA. 96-well ELISA plates were coated overnight at about 4°C with 100 µl/well M2e peptide in PBS (5 µg/ml). Plates were blocked with 200 µl/well of Assay Diluent Buffer (ADB; BD Pharmingen) for one hour at room temperature. The plates were washed three times in PBS-T. Dilutions of the sera in ADB were added (100 µl/well) and the plates were incubated overnight at about 4°C. The plates were washed 3x with PBS-T. Bound IgG was detected using HRP-conjugated goat anti-rabbit IgG (Jackson Immunochemical). The plates were washed three times with PBS-T. After adding TMB Ultra substrate (Pierce) and monitoring color development, O.D. 450 was measured on a Molecular Devices Spectramax microspectrophotometer. Results are reported as the Delta O.D. which is determined by subtracting the O.D. 450 reading for the prebleed of each animal from the O.D. 450 for each animal post-immunization.

BALB/C MOUSE EFFICACY MODEL

In a typical experiment, about 5-6 week old female BALB/c mice (10-20 per group) were obtained and allowed to acclimate for one week. Fusion proteins formulated in PBS or other suitable formulation were administered by s.c. injection. Mice were immunized on days 0 and 14. On day 21, sera was harvested by retro-orbital puncture and evaluated for M2e specific IgG by ELISA. Mice were challenged by intranasal administration of 1xLD90 of the well characterized mouse adapted Influenza A strain, A/Puerto Rico/8/34 (H1N1). Mice were monitored daily for 14 days for survival and weight loss. Mice that lost about 30% of their initial body weight were humanely sacrificed, and the day of sacrifice recorded as the day of death. Efficacy data were reported as survival times.

RESULTS

IN VITRO BIOACTIVITY

These assays were based on cell lines expressing the relevant TLR and screened for the ability to produce either IL8 or TNF-α in response to TLR triggering. In Figure 52, the ability of STF2.4xM2e (■) or STF2.OVA(○) to stimulate TLR5 dependent IL8 production was evaluated following the stimulation of TLR5 positive, HEK293 cells. The results indicate that both fusion proteins stimulated IL8 production in a dose dependent manner and that the activity of the PAMP was retained in the context of the fusion.

TLR2 activity was similarly evaluated for Pam3Cys.M2e following stimulation of TLR2 positive RAW264.7cells. In Figure 51, the experimental groups are: the known endotoxin, LPS, as a positive control (♦), LPS plus the inhibitor of endotoxin polymixin B (PMB) as a negative control (○), free Pam3Cys as a positive control for TLR2 signalling (■), free Pam3Cys plus PMB (□), Pam3Cys.M2e (♦) and Pam3Cys.M2e plus PMB (◇). The results showed similar activity profiles for Pam3Cys.M2e and the free TLR2 ligand Pam3Cys. The addition of polymyxin B (PMB) did not reduce its activity, indicating that there is no or low endotoxin contamination.

PHYSICAL LINKAGE OF PAMP AND ANTIGEN ENHANCES IMMUNOGENICITY

Using mouse models of immunogenicity, chemical coupling of Pam3Cys to M2e enhances the immunogenicity of the M2e antigen as compared to either the M2e peptide delivered alone or the M2e peptide co-delivered with free Pam3Cys. In the experiment shown in Figure 52, groups of mice were immunized on days 0 and 21 with PBS as a negative control (*), the free TLR2 ligand, Pam3CSK-4 (), M2e peptide alone (○), free Pam3CSK-4 mixed with M2e peptide (□), or the fusion of Pam3Cys and M2e referred to as Pam3.M2e (♦). The relevant the molar ratio of M2e peptide delivered was held constant. On day 28, sera were harvested and analyzed for M2e-specific antibody titers by ELISA. The results show that chemical coupling of Pam3Cys to the M2e (Pam3Cys.M2e) generates a detectable serum antibody response to the M2e antigen.

Physical linkage between the TLR5 ligand STF2 and antigen was demonstrated using the model antigen ovalbumin (OVA). Mice received a single s.c. immunization with STF2, OVA, STF2.OVA fusion protein, STF2 + OVA mixture or PBS alone. Dosages were calculated to deliver 12 μg equivalents of STF2 and OVA per group. Seven days later, sera were harvested and OVA-specific antibodies were examined by ELISA. Data shown in Figure 53 depict IgG1 titers at a 1:100 dilution of the sera. These results demonstrate that physical linkage of the TLR5 ligand and antigen results in optimal immunogenicity in vivo.

PAMP LINKED ANTIGENS ARE MORE IMMUNOGENIC THAN CONVENTIONAL ADJUVANT

Groups of 5 BALB/c mice were immunized on day 0 and 14 with 30 μg of Pam3Cys.M2e (♦), 22.5 μg of M2e which is the molar equivalent of M2e in 30 μg of Pam3Cys.M2e (◇), 22.5 mg of M2e adsorbed to the conventional adjuvant Alum (□), or 25 mg of the recombinant protein STF2.4xM2e (■). A group receiving PBS was included as a negative control (○). Sera were harvested 7 days post the second dose and M2e specific IgG were evaluated by ELISA. The results shown in Figure 54 indicate that M2e alone is poorly immunogenic in that it failed to elicit antibody titers above background. The conventional adjuvant Alum provided a modest enhancement in the immune response to M2e. The PAMP linked M2e constructs; however, provided the greatest enhancement in immunogenicity. These results indicate direct linkage of PAMPs with portions of an integral membrane protein of an influenza viral protein can elicit immune responses that are more potent than those elicited by the conventional adjuvant Alum.

DOSE AND IMMUNOGENICITY

Dose ranging studies were carried out to further assess the potency of Pam3Cys.M2e and STF2.4xM2e. For STF2.4xM2e, BALB/c mice were immunized on day 0 and 14 with dilutions of STF2.4xM2e that ranged from 0.25 to 25 μg of STF2.4xM2e per immunization. The prefix D002 refers to the specific batch of STF2.4xM2e used in this experiment, while R-028 refers to a historical reference batch of STF2.4xM2e used in this experiment. Seven days following the last immunization (Day 21) mice were bled and M2e-specific IgG responses were

evaluated by ELISA. The results shown in Figure 55 demonstrate that immunization with doses as low as 0.25 μg per immunization of STF2.4xM2e induced detectable levels of M2e-specific IgG, with the optimal dose in mice falling in the range of about 2.5 to about 25 μg.

For Pam3Cys.M2e, BALB/c mice were immunized on day 0 and 14 with 0.05 to 30 μg of Pam3Cys.M2e per immunization. Seven days following the last immunization (Day 21) mice were bled and M2e-specific IgG responses were evaluated by ELISA. The results shown in Figure 56 demonstrate that immunization with concentrations as low as 0.05 μg of Pam3Cys.M2e induced detectable levels of M2e-specific IgG, with the optimal dose for mice in this study of about 30 μg.

IMMUNOGENICITY IN MULTIPLE MOUSE STRAINS

The immunogenicity of Pam3Cys.M2e was evaluated in multiple mouse strains including BALB/c (●), C57BL/6 (■), CB6/F1 (◆), DBA/2 (▲), Cr:NIH (Swiss) (X) and C3H/HeN (*). Groups of five for each strain were immunized on day 0 and 14 with 30 μg of Pam3Cys.M2e per immunization. Sera were harvested on day 21 and levels of M2e-specific IgG evaluated by ELISA. All strains exhibited significant levels of M2e-specific IgG indicating that the immunogenicity of Pam3Cys.M2e is not dependent on a particular MHC (Figure 57).

IMMUNOGENICITY IN RABBITS

Studies aimed at evaluating the immunogenicity of Pam3Cys.M2e and STF2.4xM2e in a second species, rabbit, were carried out. In the first study, rabbits (3 females and 3 males/group) were immunized with 500, 150, 50, 15 or 5 μg (i.m.) of Pam3Cys.M2e on day 0, 21 and 42. As a control, an additional group received the formulation buffer F111 (10 mM Tris, 10 mM histidine, 75 mM NaCl, 5% sucrose, 0.02% Polysorbate-80, 0.1 mM EDTA, 0.5% ethanol, 20 mg/mL hydroxypropyl-beta-cyclodextrin, pH 7.2). On day 7 post-boost 2, peripheral blood was obtained and the anti-M2e antibody titers were evaluated by ELISA. The results shown in Figure 58 depict the individual rabbit antibody titers at a 1:125 dilution of the sera. The data suggest a dose-response relationship between the

amount of Pam3Cys.M2e used for prime/boost vaccinations and the level of the antibody titer achieved.

In the second study, rabbits (3 females and 3 males/group) were immunized with 500, 150, 50, 15 or 5 μg (i.m.) of STF2.4xM2e. As a control, an additional group received saline alone. On day 14 post-immunization, peripheral blood was obtained and the anti-M2e antibody titers were evaluated by ELISA. Notably, significant M2e-specific IgG responses were detectable by day 14 post-prime in all animals immunized (Figure 59). The results indicate that STF2.4xM2e elicits a rapid and consistent immune response in rabbits.

EFFICACY IN THE MOUSE CHALLENGE MODEL

The efficacy of the Pam3Cys.M2e and STF2.4xM2e was evaluated in BALB/c mice using the well characterized mouse adapted strain, Influenza A/Puerto Rico/8/34 (PR/8) as the challenge virus. Groups of ten mice were immunized s.c. on day 0 and 14 with 30 μg of Pam3Cys.M2e in the formulation buffer F111 (■), 30 μg of Pam3Cys.M2e in the proprietary buffer F120 (10 mM Tris, 10 mM histidine, 10% sucrose, 0.02% Polysorbate-80, 0.1 mM EDTA, 0.5% ethanol, 0.075% docusate sodium, pH 7.2) (▲), 30 μg of Pam3Cys.M2e in the buffer F119 (10 mM Tris, 10 mM histidine, 75 mM NaCl, 5% sucrose, 0.02% Polysorbate-80, 0.1 mM EDTA, 0.5% ethanol, 0.1% docusate sodium, pH 7.2), 30 μg of STF2.4xM2e in the buffer F105 (10 mM Tris, 10 mM histidine, 75 mM NaCl, 5% sucrose, 0.02% Polysorbate-80, 0.1 mM EDTA, 0.5% ethanol, pH 7.2), 3 μg of STF2.4xM2e in buffer F105 (10 mM Tris, 10 mM histidine, 75 mM NaCl, 5% sucrose, 0.02% Polysorbate-80, 0.1 mM EDTA, 0.5% ethanol, pH 7.2) (●) or 0.3 μg of STF2.4xM2e in buffer F105 (□). A group receiving PBS alone was included as a negative control (○), and a convalescent group with immunity to PR/8 following a sublethal challenge with the virus was included as a positive control (◇). On day 28, animals were challenge with an LD90 of the PR/8 challenge stock. Weight loss and survival was followed for 14 days post challenge (Figure 60).

Animals in the convalescent group which had successfully cleared an earlier non-lethal infection with PR/8 demonstrated 100% protection to a subsequent viral challenge. Animals receiving the PBS buffer alone exhibited morbidity beginning

on days 7 and 8, with 80% lethality occurring by day 10, while animals immunized with 30 μg of Pam3Cys.M2e in F111 demonstrated enhanced survival, with 50% of mice surviving the challenge. Animals receiving Pam3Cys.M2e in F119 exhibited morbidity beginning on days 8 and 9 with 80% of the mice surviving. Animals receiving Pam3Cys.M2e in buffer F120 (10 mM Tris, 10 mM histidine, 10% sucrose, 0.02% Polysorbate-80, 0.1 mM EDTA, 0.5% ethanol, 0.075% docusate sodium, pH 7.2) or the STF2.4xM2e protein exhibited the mildest disease course with 90 to 100% of the mice in these groups surviving the lethal challenge. These results demonstrate that both Pam3Cys.M2e and STF2.4xM2e can confer protective immunity to a challenge with influenza A in vivo.

DISCUSSION

Salmonella typhimurium flagellin (fljB) is a ligand for TLR5. A recombinant protein consisting of full-length flijB (STF2) fused to four tandem repeats of M2e was expressed in *E. coli* and purified to > 95% purity with low endotoxin levels. In reporter cell lines, this protein (STF2.4xM2e) triggered IL8 production in a TLR5-dependent fashion. Mice immunized with dilutions of STF2.4xM2e that ranged from 0.25 μg to 25 μg, formulated in the buffer F105 which is without a conventional adjuvant or carrier, mounted a vigorous antibody response. The potency of the recombinant protein was further demonstrated in rabbit immunogenicity studies where animals receiving as little as 5 μg of protein seroconverted after a single dose. The efficacy of the PAMP fusion protein was demonstrated in the mouse challenge model using Influenza A/Puerto Rico/8/34 as the challenge virus. Mice immunized with as little as about 0.3 μg of the protein per dose exhibited mild morbidity with 100% of the mice surviving the challenge.

Synthetic tripalmitoylated peptides mimic the acylated amino terminus of lipidated bacterial proteins and are potent activators of TLR2. In these studies, a tripalmitoylated peptide consisting of three fatty acid chains linked to a cysteine residue and the amino terminus of the Influenza A M2 ectodomain (M2e) was synthesized using standard solid-phase peptide chemistries. This peptide (Pam3Cys.M2e) triggered TNFα production in a TLR2-dependent fashion in reporter cell lines. When used to immunize mice without adjuvant, Pam3Cys.M2e

generated an antibody response that was more potent than M2e when mixed with free Pam3CSK-4. Pam3Cys.M2e was also found to be immunogenic in rabbits where a dose response relationship was observed between the amount of Pam3Cys.M2e used for immunization and the antibody titer achieved. The efficacy of the Pam3Cys.M2e peptide in a number of different formulations was evaluated in the mouse challenge model using Influenza A/Puerto Rico/8/34 as the challenge virus. Pam3Cys.M2e formulated in F119 and F120 exhibited the mildest morbidity with about 80 to about 100% of the mice surviving the challenge.

Example 20

Materials and Methods

PCR Amplification and DNA primers

All PCR amplifications were performed using Pfu Ultra Hotstart PCR Master Mix (Catalog number 600630) from Stratagene (La Jolla, CA) according to the manufacturer's recommendations. DNA primers were purchased from Sigma Genosys and are described below.

STF28BGF-1:
CTCGGGAGATCTGCACAAGTAATCAACACTAACAGTCT (SEQ ID NO: 644)

STF28MCR-1:
CCATGGGCTAGCAGGATCCACCGGCGCTCCCTGCACGTTCA (SEQ ID NO: 645)

STF28MCF-2:
GGAGCGCCGGTGGATCCTGCTAGCCCATGGACCGAAAACCCG (SEQ ID NO: 646)

STF28ECR-2:
TCTGCAGAATTCACGTAACAGAGACAGCACGTTCTGCGGGACGTCCCGC AGAACGTGCTGTCTCTGTTACGTGAATTCTGCAGA (SEQ ID NO: 647)

pET24AR:5 TCCGGCGTAGAGGATCGAGA (SEQ ID NO: 648)

STF2-E3R3:
CAATTGACCTTCAAGCTTCGAATTGCCCTTACGTAACAGAGACAGCACGT TCTG (SEQ ID NO: 649)

AX-E3F3:
AAGCTTGAAGGTCAATTGGAATTCCCTAGGACTAGTATGGAAAAATTGC AGTTGAAG (SEQ ID NO: 650)

pET24AF: GCTTAATGCGCCGCTACAGG (SEQ ID NO: 651)

5'WNE28: GCGGCCGCTCATGGAAAAATTGCAGTTGAAGGGAACAACC (SEQ ID NO: 652)

3'WNE28: CCGCGGTTTGCCAATGCTGCTTCCAGACTTGT (SEQ ID NO: 653)

NdeI-STF2: CCGGCATGCCATATGGCACAAGTAATCAACACTAACAGTCTGTCGCTGC (SEQ ID NO: 654)

BlpI-EdIII: GCATGCTCAGCTTATTAAGGGTTTGCCAATGCTGCTTCCCAGACTTGTG (SEQ ID NO: 656)

JE EIII primer: TACGTGAATTCAGCAGATATCCAGCAC (SEQ ID NO: 53)

Cloning of pET/STF2Δ.EIII

Full length flagellin of *Salmonella typhimurium* fljb (flagellin phase 2) (also referred to herein as "STF2") is encoded by a 1.5 kb gene. A truncated version of the STF2 (STF2Δ, SEQ ID NO: 606, encoded by SEQ ID NO: 607) was generated by deleting the hyper-variable region that spans amino acids 170 to 415 of SEQ ID NO: 604. The deleted region was replaced with a short flexible linker (GAPVDPASPW, SEQ ID NO: 659) designed to facilitate interactions of the NH2 and COOH termini sequences necessary for TLR5 signaling. To generate this construct, a two-step PCR was used. In the first reaction, STF2.OVA (SEQ ID NO: 755 encoding amino acid sequence SEQ ID NO: 756) served as the DNA template and STF28BGF-1 and STF28MCR-1 were used as primer pairs. In a separate reaction, the same DNA template was combined with primers STF28MCF-2 and STF28ECR-2.

The PCR amplification reactions generated about 500 bp and about 270 bp fragments, respectively. These PCR products were combined in a final PCR reaction using STF28BGF-1 and STF28ECR-2 as primers. The amplified DNA product from this reaction (about 0.77 kb) was digested with *Bgl*II and *EcoR*I restriction enzymes and ligated into pMTBiP/V5-His B (Invitrogen, Carlsbad, CA) that had previously been digested with *Bgl*II and *EcoRI* and treated with alkaline phosphatase. An aliquot of the ligation mix was used to transform TOP10 cells

(InVitrogen, Carlsbad, CA). PCR screening was performed using vector specific primers, pMTFOR (methionine promoter) (CATCTCAGTGCAACTAAA, SEQ ID NO: 759) and BGHREV (bovine growth hormone poly A) (TAGAAGGCACAGTCGAGG, SEQ ID NO: 760), to identify several positive clones. All positive clones were further analyzed by restriction mapping analysis and confirmed by DNA sequencing. The resultant construct pMT/STF2Δ was used to generate pMT/STF2Δ.EIII+ .

The domain III of the West Nile virus envelope protein of pET/STF2Δ.EIII+ (SEQ ID NOS: 673, 674) was derived from the *Drosophila* expression plasmid pMT/STF2.E. This plasmid contains full-length STF2 (amino acids 1-506, SEQ ID NO: 604) fused to the West Nile Virus envelope protein (amino acids 1-406 , SEQ ID NO: 642). The pMT/STF2.E (SEQ ID NO: 761) clone AX-1 was used as a DNA template and 5'WNE28 (SEQ ID NO: 652) and 3'WNE28 (SEQ ID NO: 653) served as primers for PCR amplification. In order to facilitate restriction analysis and subsequent cloning steps, the 5' primer encoded a novel *Not*I site (New England Biolabls, Beverly, MA) and the 3' primer contained a unique *Sac*II site. The amplified EIII+ DNA fragment (345bp; SEQ ID NO: 781 that encodes amino acids 292-406 of SEQ ID NO: 642) was subcloned into pCR-Blunt II-TOPO cloning vector (InVitrogen, Carlsbad, CA) to generate plasmid TOPOEIII. A stop codon was subsequently introduced downstream of the EIII+ sequence by blunting the *Sac*II and *Spe*I restriction sites using T4 DNA polymerase.

To generate pMT/STF2Δ.EIII+ (SEQ ID NOS: 673, 674), the EIII+ fragment was isolated from TOPOEIII+ using *Not*I and *BamH*I restriction sites and ligated into the *Not*I and *Sac*II restriction sites in pMT/STF2Δ. The *BamH*I site of the EIII+ DNA fragment and the *Sac*II site of pMTSTF2Δ were blunted with T4 DNA polymerase prior to ligation. The STF2Δ.EIII+ sequence (SEQ ID NOS: 673, 674) from pMT/STF2Δ.EIII+ was isolated by PCR amplification using the primers NdeI-STF2 and BlpI-EdIII. To generate pET/STF2Δ.EIII+ (SEQ ID NO: 674), the PCR product was digested with *Nde*I and *Blp*I and ligated into pET24a plasmid that had been predigested with *Nde*I and *Blp*I. The ligation mix was transformed into Mach-1 cells (InVitrogen, Carlsbad, CA) and the cells were grown on LB supplemented

with 50 μg/ml kanamycin. Several colonies were screened by restriction mapping and were verified by DNA sequencing.

Cloning of pET/STF2.EIII+

The West Nile virus EIII+ sequence of pET/STF2.EIII+ (SEQ ID NOS: 657, 658) was derived from pETSTF2.E (SEQ ID NOS: 761, 762). This *E. coli* expression plasmid contains full-length STF2 (amino acids 1-506) fused to the West Nile Virus envelope protein (amino acids 292-406 of SEQ ID NO: 642, which is SEQ ID NO: 610). In two independent PCR reactions, pET/STF2.E was used as the DNA template. One reaction used the primers pET24AR:5 (SEQ ID NO: 648) and STF2-E3R3: (SEQ ID NO: 649) and the other used AX-E3F3 (SEQ ID NO: 650) and pET24AF (SEQ ID NO: 651). These PCR reactions generated a 1.5 kb fragment that consisted of full-length STF2 and a 340bp fragment that comprised the EIII domain plus additional amino acids that extended into domain I of the envelope protein. Aliquots of these PCR amplification reactions were combined, and the two products served as templates for a PCR reaction with the external primers pET24AR (SEQ ID NO: 648) and pET24AF (SEQ ID NO: 651). This resulted in the generation of about a 1.8kb DNA fragment that fused EIII+ sequence (SEQ ID NO: 781, a nucleic acid sequence encoding amino acids 292-406 of SEQ ID NO: 642, which is SEQ ID NO: 610) to STF2. The PCR product was digested with *Nde*I and *Blp*I and gel purified and ligated by compatible ends to a pET24a vector that had previously been digested with compatible enzymes and de-phosphorylated. The ligation mix was transformed into Mach-1 cells (InVitrogen, Carlsbad, CA) as described for pET/STF2Δ.EIII+. Several colonies were screened by restriction mapping and two clones were verified by DNA sequencing.

Cloning of pET/STF2Δ.JEIII+

A portion of the envelope protein of a *Japanese encephalitis* virus (JEV) (strain SA-14-14-2 (Jai, L., *et al., Chin Med J (Eng) 116*:941-943 (2003)); currently employed in a JEV vaccine encoded by domain III was custom synthesized by DNA 2. Inc (Menlo Park, CA). The portion of domain III was excised from the pJ2:G01510 using NotI and Blp I site that flank the insert. The DNA insert was gel isolated and cloned by compatible ends to pET24A/STF2Δ.EIII+ (SEQ ID NOS:

673, 674) that had previously been digested with the appropriate enzymes to release the West Nile virus EIII+ insert. The deleted vector was then gel purified and ligated to an aliquot of JE EIII+. The ligation mix was used to transform TOP-10 cells (InVitrogen, Carlsbad, CA) and the cells were grown on LB supplemented with 50 μg/ml kanamycin. Several colonies were screened by restriction mapping and were verified by DNA sequencing.

The resulting construct, pET24A/STF2Δ.JEIII (SEQ ID NOS: 608, 609) was BLR (DE3) strain (Novagen) and expression was monitored in several clones using Commassie Blue staining which was confirmed by Western blot using anti-flagellin antibodies. Using, pET24A/STF2Δ.JEIII+ as the DNA template and the JE EIII+ oligonucleotide as primer (SEQ ID NO: 656) the cysteine residue in the linker region between STF2Δ and JEIII+ was changed to a serine residue using QuikChange Site Directed Mutagenesis Kit (Stratagene, LaJolla, CA) according to the manaufacturer's instructions. The clone was verified by sequencing and assayed for expression as described for pET24A/STF2Δ.JEIII+ above.

When a cysteine residue in a linker in change to a serine residue the fusion protein in also referred to herein by includsion of an "s" in the designation of the fusion protein. For example, "STF2Δ.EIII+" includes a cysteine residue in the linker (SEQ ID NO: 674), whereas "STF2Δ.EIIIs+" include a serine residue substituted for the cysteine residue in the linker (SEQ ID NO: 675).

Cloning the EIII domain of each dengue virus fused to the C-terminal end of flagellin (STF2Δ)

Initially, obtaining biologically active material from the fusion of the entire envelope protein of West Nile virus was difficult, perhaps due to the presence of multiple cysteines residues (12 cysteines) in the envelope protein (see SEQ ID NO: 642). However, when the region encoding domain III (EIII) of the protein was sub-cloned, the fusion protein was abundantly expressed *in E. coli* and was highly efficacious in mice. Although there is an overall sequence dissimilarity among the 4 distinct DEN viruses (Den1, Den2, Den3, Den4, SEQ ID NOS: 763-770, the three-dimensional structures within domain III of the envelope protein are similar among the flaviviruses. This domain in DEN and other flaviviruses encodes the majority of the type-specific contiguous critical/dominant neutralizing epitopes. Domain III of

the dengue viruses (Den1, Den2, Den3 and Den4) has been expressed in bacteria and shown to be immunogenic, capable of inducing neutralizing antibodies in experimental animals (Simmons, M., *et al., Am. J. Trop. Med Hyg 65*:159 (2001)). Domain III corresponding to residues about 295 to about 399 (exact numbering depends on the particular DEN virus, for example, of SEQ ID NOS: 763, 765, 767, 769) of the four different DEN viruses have been codon-optimized for expression in *E. coli*. The synthetic gene was amplified by using PCR and sub-cloned into the *Not*I site of the vector pET/STF2Δ generating pET/STF2Δ.DEN1EIII, pET/STF2Δ.DEN2EIII, pET/STF2Δ.DEN3EIII and pET/STF2Δ.DEN4EIII (SEQ ID NOS: 683, 685, 687 and 689).

E. coli production of STF2.EIII+, STF2Δ.EIII+, STF2Δ.EIIIs+ and STF2Δ.JEIII+

Cell cultures (6L) of BLR(DE3) pLysS that harbor pETSTF2.EIII+ (SEQ ID NOS: 657, 658), pETSTF2Δ.EIII+ (SEQ ID NOS: 673, 674), pETSTF2Δ.EIIIs+ (SEQ ID NOS: 675, 676) or pETSTF2Δ.JEIII+ SEQ ID NOS: 608, 609) were grown in LB medium containing 15 μg/ml kanamycin, 12.5 μg/ml tetracycline and 24 μg/ml chloramphenicol. At an OD_{600} of about 0.6 protein expression was induced with 1 mM IPTG for about 3 h at about 37°C. Following induction, cells were harvested by centrifugation (7000 rpm x 7 minutes in a Sorvall RC5C centrifuge) and resuspended in 2x PBS, 1% glycerol, DNAse, 1 mM PMSF, protease inhibitor cocktail and 1 mg/ml lysozyme. The suspension was passed through a microfluidizer to lyse the cells and the lysate was centrifuged (45,000 g for one hour in a Beckman Optima L ultracentrifuge) to separate the soluble fraction from inclusion bodies. Under these growth and induction conditions, STF2.EIII+ was expressed as a soluble protein and STF2Δ.EIII+ (SEQ ID NOS: 673, 674), STF2Δ.EIIIs+ (SEQ ID NOS: 675, 676) and STF2Δ.JEIII+ (SEQ ID NOS: 608, 609) formed inclusion bodies.

Purification of STF2.EIII+

Cell lysate containing soluble STF2.EIII+ (SEQ ID NOS: 657, 658) was applied to Sepharose Q resin (Amersham Biosciences, Piscataway, NJ) in the presence of 0.5 M NaCl to reduce DNA, endotoxin, and other contaminants. The flow-through fraction was collected and the conductivity adjusted by a 10-fold

dilution with buffer A (Buffer A: 100 mM Tris-Cl, pH 8.0). The diluted material was re-loaded onto Q Sepharose and bound protein was eluted with a linear gradient from 20% to 60% Buffer B (Buffer B: 100 mM Tris-Cl, 1 M NaCl, pH 8.0). Fractions containing STF2.EIII+ were pooled and further processed by Superdex-200 gel (SD200) filtration chromatography in the presence of Na-deoxycholate to remove residual endotoxin (running buffer: 1% Na-deoxycholate, 100 mM NaCl, 100 mM Tris-HCl, 1% glycerol, pH 8.0). Following SD200 chromatography, the eluted protein was loaded directly onto Q Sepharose and washed extensively with buffer A to remove detergent. Bound protein was again eluted with a linear gradient from 20% to 60% Buffer B. In one preparation (Batch 057), this step was substituted with a detergent removal procedure using Extract-D detergent removal gel (Pierce Biotechnology, Rockford, IL). The purified protein was dialyzed against buffer containing 50 mM Tris, 100 mM NaCl and 1% glycerol and stored at -80°C.

Purification of STF2Δ.EIII+

STF2Δ.EIII+ inclusion bodies were collected by low-speed centrifugation (7000 rpm x 7 minutes in a Sorvall RC5C centrifuge) and solubilized with buffer containing 8 M urea, 100mM Tris-HCl, 5 mM EDTA, pH 8.0. The urea concentration of the solubilized protein was adjusted to 1 M and the sample was loaded onto Q Sepharose. The bound protein was eluted using a linear gradient from 0% to 100% Buffer B. (Buffer A: 100 mM Tris-HCl, 5mM EDTA, 1 M urea, pH 8.0. Buffer B: 100 mM Tris-Cl, 5 mM EDTA, 1 M NaCl, 1 M urea, pH 8.0). Due to the formation of protein aggregates following elution, the urea concentration of the Q Sepharose material was adjusted to 8 M. The protein was further purified by gel filtration chromatography using SD200. The column was pre-equilibrated with 100 mM Tris-HCl, pH 8.0, 100 mM NaCl, 1% glycerol, 8 M urea plus 1% Na-deoxycholate. The eluted protein was subjected to a second IEX chromatography step using Source Q to remove 1% Na-deoxycholate. Bound protein was eluted with a linear gradient from 20% to 60% Buffer B. (Buffer A: 100 mM Tris-Cl, pH 8.0, 8 M urea, 5mM EDTA. Buffer B: 100 mM Tris-HCl, pH 8.0, 5 mM EDTA, 8 M urea, 1 M NaCl). Final polishing of the protein was completed by gel filtration chromatography using SD200 (Running Buffer: 100 mM Tris-HCl, pH 8.0, 8 M

urea, 100 mM NaCl and 1% glycerol). Reducing agent was added to the SD200 fraction (2.5 mM DTT) and the protein was refolded by step-wise dialysis against decreasing concentrations of urea. The urea concentration was reduced sequentially against buffers that contained 100 mM Tris-HCl, pH 8.0, 100 mM NaCl, 1% glycerol and 6 M, 4 M, 2 M or no urea.

Refolding and purification of STF2Δ.EIII+ trimer

STF2Δ.EIII+ (SEQ ID NOS: 673, 674) from urea-solubilized inclusion bodies was efficiently refolded to form trimer product by simple dialysis as described above the trimer (3 of thes STFΔ.EIII fusion proteins) was deduced based on molecular weight in SDS-PAGE. Following dialysis, endotoxin was removed by multiple extractions with Triton X-114. The trimer was purified and separated from monomer and aggregates by S200 size exclusion chromatography. The final product migrated as a single band with an apparent molecular weight of about 130 kDa on SDS-PAGE.

Refolding and purification of STF2Δ.EIII+ monomer

The monomeric form of STF2Δ.EIII+ (SEQ ID NOS: 673, 674) was produced consistently and efficiently by refolding using rapid dilution, which prevented individual STF2Δ.EIII+ fusion proteins from interacting with one another to form meutimers, such as trimers (*supra*). STF2Δ.EIII+ solubilized from inclusion bodies in 4M urea was raised to 8M urea without reductant. The protein was then rapidly diluted in Tris/NaCl/glycerol buffer, pH 8.0, to about 0.1mg/ml and a final urea concentration of 0.1M at room temperature. The monomer was further purified and separated from aggregates by S200 size exclusion chromatography. The final product migrated as a single band with an apparent molecular weight of about 43 kDa on SDS-PAGE.

Purification of STF2Δ.EIIIs+ (serine substitution of the linker between STF2Δ and EIII+, SEQ ID NO: 675)

STF2Δ.EIIIs+ (SEQ ID NOS: 675, 676) from solubilized inclusion bodies was refolded using a rapid dilution method similar to that used to refold the STF2Δ.EIII+ monomer. The refolded protein was captured on a butyl sepharose

column and eluted while removing most of the endotoxin contamination. Eluate from the butyl sepharose purification was concentrated and put through 4 cycles of Triton X-114 extractions to reduce endotoxin levels down to about <0.1 EU/μg before a final purification step over SD200 gel filtration. The final pooled product migrated as a single band with an apparent molecular weight of about 43 kDa on SDS-PAGE and contained a trace amount of Triton X-114 (about 0.000015%).

Purification of STF2Δ.JEIII+ (SEQ ID NOS: 608, 609)

Protein was isolated from inclusion bodies under denaturing conditions. Inclusion bodies were washed with detergent (0.5% Triton X 100) and solubilized in 8 M urea, resulting in partial purification of the target protein. For endotoxin removal, protein was applied on a Source S cation exchange column at low pH (about 3.5) and eluted with a salt gradient (0 to about 1M NaCl). The protein was refolded using rapid dilution as described for STF2Δ.EIII+ monomer. The protein was then concentrated and further purified using SD200 to separate the monomeric form of the protein from aggregates. The purified material migrated with an apparent molecular weight of about 43 kDa on SDS PAGE and contained acceptable levels of endotoxin (about 0.03 EU/ug).

Fed Batch Production of Fusion Proteins

STF2Δ.EIIIs+ was produced in an aerobic bioreactor using a fed batch process. Three control loops were placed to control pH by acid (2 N HCl) or base (3 N NH_4OH) addition, temperature by heating (heating blanket) or cooling (time cycled cooling loop), and dissolved oxygen by compressed air flow (manually controlled), agitation (mixing speed) and O_2 flow (timed cycled) in cascade. Cells [BLR(DE3) pLysS that harbor the STF2Δ.EIIIs+ were adapted to and banked in MRSF media (see *infra*), and frozen in 25% glycerol. Cells were scaled up for the bioreactor by adding 1 mL of banked cells to 1 L of MRSF media and agitating at about 37°C for about 15.5 to about 16.5 hours. Cells from the scale up process were added in a about 1:10 ratio to MRSF or MRBR synthetic media at about 37°C and about 0.5 vvm air flow.

The process was run in batch mode at about 37°C until the cells oxygen consumption was such that the compressed air flow is about 1.5 vvm and the

agitation is at the maximum, about 6 hours, when the temperature is dropped to between about 25°C and about 33°C. The feed can be started before the culture is induced, or up to about 1 and about 1/2 hours after. The feed rate can be kept constant, or adjusted based on process variables (dissolved oxygen, glucose concentration). The culture was induced with IPTG upon batch glucose exhaustion. The culture was maintained for a minimum of about 2 hours and a maximum of about 20 hours.

MRBR Media Composition	g/L
Glucose	20
KH_2PO_4	2.2
$(NH_4)_2SO_4$	4.5
Citric Acid	1.0
$MgSO_4(7H_20)$	1.0
$CaCl_2$	0.04
Trace Metals	1ml
Thiamine HCl	0.01
Antifoam	0.05

Trace Metal Solution 1000x Component	g/L
EDTA, disodium	5
$FeSO_4(7H_2O)$	10
$ZnSO_4(7H_2O)$	2
$MnSO_4(H_2O)$	2
$CoCl_2(6H_2O)$	0.2
$CuSO_4(5H_2O)$	0.1
$Na_2MoO_4(2H_2O)$	0.2
H_3BO_3	0.1

MRSF Media Composition	g/L
Glucose	10 (20 in bioreactor)
KH_2PO_4	7.8
$(NH_4)_2SO4$	2.33
Citric Acid	1.0
$MgSO_4(7H_20)$	1.0
$CaCl_2$	0.04
Trace Metals	1ml
Thiamine HCl	0.01
Kanamycin	0.0075 (shake flask only)

Feed Media Composition	g/L
NaCl	0.5
$FeSO_4(7H_2O)$	2
$CaCl_2$	3.5
$MgSO_4(7H_2O)$	12
Thiamine HCl	1
Trace Metals	1ml
Glucose	100

STF2Δ.EIIIs+ was produced as inclusion bodies. Upon harvest, the cells were separated from the conditioned media by centrifugation (Beckman Avanti J-20 XP, JLA 8.1000 rotor, 10 kxg for about 20 minutes at about 4°C) and resuspended in equal volume of 50 mM Tris, 100 mM NaCl, 1 mM EDTA, pH 8.0. The centrifugation was repeated under the same conditions, with the cells resuspended in a minimum volume of the same buffer. The suspension was passed through a homogenizer (APV-1000) at > 10,000 psi for at least two passes.

The solids can be separated and the STF2Δ.EIIIs solubilized by one of three methods; centrifugation, filtration, or fluidized bed chromatography.

Method 1

Solids are separated by centrifugation (Beckman Avanti J-20 XP, JA 20 rotor, 20 kxg for 20 minutes at 4°C) and resuspended in 50 mM tris, 1 m NaCl, 1mM EDTA, 1% glycerol, 0.5% Triton X-100, pH 8.0. This process was repeated up to 6 times (total) at increasing speeds and times (to a maximum of about 40 kxg for about 20 minutes). After the final pellet recovery, the pellet was resuspended in 50 mM Tris, 0.1M NaCl, 1 mM EDTA, pH 8.0 and clarified by centrifugation (Beckman Avanti J-20 XP, JA 20 rotor, 40 kxg for about 20 minutes at about 4°C) The pellet was resuspended and dissolved in 50 mM Tris, 0.1M NaCl, 1 mM EDTA, 4 M urea, pH 8.0. Insolubles were removed by centrifugation (Beckman Avanti J-20 XP, JA 20 rotor, 40 kxg for about 50 minutes at about 4°C) , the supernatant retained for further processing.

After the multiple washes described above, STF2Δ.EIIIs can also be dissolved in 50 mM acetate, 10 mM NaCl, 8M urea, pH about 4.1 to about 5.3 and clarified by centrifugation (Beckman Avanti J-20 XP, JA 20 rotor, 20 kxg for about 20 minutes).

Method 2

After homogenization, the lysate was captured in body feed and STF2Δ.EIIIs+ extracted with urea containing buffer. Body feed is a filter aid designed to trap particles in a cake above a depth filter. The body feed (Advanced Minerals Corporation CelPure 65) is a diatomite (silica powder) with a high surface area and low permeability, retaining < 0.2μm particles. The filter aid was pre-mixed with the lysate and pumped over a depth filter (Ertel 703), building up a cake containing both body feed and lysate particles. The suspension creates a depth filter as the particles settle on the filter pad. A 50mM Tris, 100mM NaCl pH 8.0 wash was performed to remove soluble proteins and nucleic acids. A subsequent wash with 50 mM Tris, 100 mM NaCl, 4 M urea, pH 8 solubilizes and removes the STF2Δ.EIIIs from the body feed for further processing.

Method 3

After the cells were initially resuspended in buffer, they were resuspended in sodium chloride and urea containing buffer at pH about 6 to about 8 and homogenized. The lysate was applied on a Streamline CST fluidized bed column (GE Healthcare) where the STF2Δ.EIIIs+ binds to the resin and the particulates flow through. STF2Δ.EIIIs+ may be eluted in low salt conditions at a pH greater than the load pH, in the presence or absence of detergents such as Triton X-100 or polysorbate 80.

SDS-PAGE

Proteins (typically about 5 μg) were diluted in SDS-PAGE sample buffer with and without β-mercaptoethanol. The samples were boiled for 5 minutes and loaded onto a 4-20% SDS polyacrylamide gel. Following electrophoresis, gels were stained with coomassie blue to visualize protein bands.

Endotoxin assay

Endotoxin levels were measured using the QCL-1000 Quantitative Chromogenic LAL test kit (BioWhittaker #50-648U, Walkersville, MD), following the manufacturer's instructions for the microplate method.

Protein Assay

Protein concentrations were determined by the MicroBCA Protein Assay Reagent Kit in a 96-well format using BSA as a standard (Pierce Biotechnology, Rockford, IL).

TLR5 bioactivity assay

HEK293 cells (ATCC, Catalog No. CRL-1573 Manassas, Virginia) constitutively express TLR5, and secrete several soluble factors, including IL-8, in response to TLR5 signaling. Cells were seeded in 96-well microplates (about 50,000 cells/well), fusion proteins added and incubated overnight. The next day, the conditioned medium was harvested, transferred to a clean 96-well microplate, and frozen at -20°C. After thawing, the conditioned medium was assayed for the presence of IL-8 in a sandwich ELISA using an anti-human IL-8 matched antibody

pair (Pierce, #M801E and #M802B, Rockford, IL) following the manufacturer's instructions. Optical density was measured using a microplate spectrophotometer (FARCyte, Amersham Biosciences, Piscataway, NJ).

Plaque reduction neturalization test (PRNT)

PRNT was performed according to Wang, *et al., J. Immunol. 167*:5273-5277 (2001). Briefly, serum samples were heat inactivated by incubation in a 56°C water bath for about 30 min and were serially diluted in PBS with 5% gelatin from 1/10 to 1/2560. West Nile virus was diluted in PBS with 5% gelatin so that the final concentration was about 100 PFU/well. Virus was mixed with about 75 µl serum in a 96-well plate at about 37°C for about 1 h. Aliquots of serum-virus mixture were inoculated onto confluent monolayers of Vero cells in a six-well tissue culture plate. The cells were incubated at about 37°C for 1 h, and the plates were shaken every 15 min. The agarose overlay was then added. The overlay was prepared by mixing equal volumes of a solution consisting of 100 ml 2x MEM (Life Technologies) with sterile 2% agarose. Both solutions were placed in a 40°C water bath for 1 h before adding the overlay. The cells were incubated for 4 days at 37°C in a humidified 5% CO_2-air mixture. A second overlay with an additional 4% neutral red was added on day 5. Virus plaques were counted about 12 h later.

Antigenicity of STF2Δ-fusion proteins

ELISA plates (96-well) were coated overnight at 4°C with serial dilutions (100 µl/well) of purified STF2Δ-fusion proteins (SEQ ID NOS: 761, 762, 657, 658, 673, 674) in PBS (about 2µg/ml). Plates were blocked with 200 µl/well of Assay Diluent Buffer (ADB; BD Pharmingen) for one hour at room temperature. The plates were washed 3x in PBS-Tween, and then incubated with antibodies reactive with flagellin or the E domain of the construct. The expression of flagellin was detected using the mAb 6H11 (Intotek), while the antigenicity of WNV-E was monitored using a panel of mAb (5C5, 7H2, 5H10, 3A3, and 3D9) (Beasley, D.W., *et al., J. Virol. 76*:13097-13100 (2002)) were purchased from Bioreliance (Road Rockville, MD). Antibodies diluted in ADB (about 100µl/well) were incubated overnight at 4°C. The plates were washed 3x with PBS-T. HRP-labeled goat anti-mouse IgG antibodies (Jackson Immunochemical, West Grove, PA) diluted in ADB

were added (100µl/well) and the plates were incubated at room temperature for 1 hour. The plates were washed 3x with PBS-T. After adding TMB (3,3',5,5'-tetramentylbenzidine) Ultra substrate (Pierce Biotechnology, Rockford, IL) and monitoring color development, A_{450} was measured on a Tecan Farcyte microspectrophotometer.

Immunization of mice

C3H/HeN mice (10 per group) were immunized intraperitoneally or subcutaneously with the indicated concentrations of fusion proteins or synthetic peptides on days 0, 14 and 28 . On days 21 and 35, immunized animals were bled by retro-orbital puncture. Sera were harvested by clotting and centrifugation of the heparin-free blood samples. On day 35, mice were challenged with a lethal dose of WNV strain 2741 (Wang, T., *et al., J. Immunol 167*:5273-5277 (2001)). Survival was monitored for 21 days post-challenge.

Serum antibody determination

West Nile envelope protein specific IgG levels were determined by ELISA. ELISA plates (96-wells) were coated overnight at about 4°C with 100 µl/well of West Nile E protein mAb 5C5, 7H2, 5H10, 3A3, and 3D9 (Beasley, D.W., *et al., J. Viro. 76:*13097-13100 (2002)) (Bioreliance, Road Rockville, MD) in PBS at a concentration of 2µg/ml. Plates were blocked with 200 µl/well of Assay Diluent Buffer (ADB; BD Pharmingen, San Diego CA) for one hour at room temperature. The plates were washed 3x in PBS-T. Dilutions of the sera in ADB were added (100µl/well) and the plates were incubated overnight at 4°C. The plates were washed 3x with PBS-T. HRP-labeled goat anti-mouse IgG antibodies (Jackson Immunochemical, West Grove, PA) diluted in ADB were added (100µl/well) and the plates were incubated at room temperature for 1 hour. The plates were washed 3x with PBS-T. After adding TMB (3,3',5,5'-tetramentylbenzidine) Ultra substrate (Pierce Biotechnology, Rockford, IL) and monitoring color development, A_{450} was measured on a Tecan Farcyte microspectrophotometer.

Production of Pam3Cys.WNV001 peptide synthesis

Pam3Cys.WNV001 was synthesized by Bachem Bioscience, Inc. (King of Prussia, PA). WNV001 is a 20 amino acid peptide (SEQ ID NO: 771) of the West Nile virus envelope protein chemically coupled to a tri-palmitoylcysteine (Pam3Cys) moiety through the amino terminal serine residue of the peptide. The chemical name for Pam3Cys.WNV001 is [Palmitoyl-Cys((RS)-2,3-di(palmitoyloxy)-propyl)-LTSGHLKCRVKMEKLQLKGT (SEQ ID NO: 771) acetate salt]. The molecular mass of Pam3Cys.WNV001 is 3163.3 daltons. The peptide was synthesized by Bachem using solid phase synthesis methodologies and FMOC chemistry. The amino acid sequence of Pam3Cys.WNV001 was assembled on an H-Pro-2-chlorotrityl chloride resin by solid phase peptide synthesis. The peptide chain was elongated by successive coupling of the amino acid derivatives. Each coupling step was preceded by an Fmoc-deprotection step and were accompanied by repeated washing of the resin. After coupling of the last amino acid derivative, the final Fmoc-deprotection step was performed. Finally, the peptide resin was washed and dried under reduced pressure. During solid phase peptide synthesis color indicator tests were performed for each step to monitor the completion of the Fmoc-cleavage and the subsequent coupling of the amino acid derivatives. To couple Pam3Cys-OH to the elongated peptide, the lipid moiety was pre-activated with N,N'-dicyclohexyl-carbodiimide (DCCI) in the presence of 1-hydroxybenzotriazole (HOBt). The resulting solution was filtered and added to the peptide resin. At the end of the reaction time the peptide resin was washed and dried under reduced pressure. Color indicator tests were performed to control the coupling of Pam3Cys-OH. The completed peptide was cleaved from the resin by incubating with trifluroacetic acid (TFA). The liberated product (crude peptide material) was precipitated from the reaction mixture and lyophilized. The crude product was used for initial immunogenicity studies.

Synthesis of WNV-E peptide arrays

Peptide arrays (SEQ ID NOS: 703-754) were synthesized by Sigma Genosys (Woodlands, TX).

RESULTS

West Nile fusion protein

West Nile virus (WNV) has emerged in recent years in temperate regions of Europe and North America, presenting a threat to public and animal health. The most serious manifestation of WNV infection is fatal encephalitis (inflammation of the brain) in humans and horses, as well as mortality in certain domestic and wild birds. WNV has also been a significant cause of human illness in the United States. The envelope glycoprotein of West Nile (WNV-E) and other flaviviruses may generate neutralizing and protective antibodies. By linking this antigen to a Toll-like Receptor ligand, the compositions, fusion proteins and polypeptides described herein may target appropriate antigen presenting cells without the need for adjuvant or other immune modulator formulations.

As described herein, several strategies have been implemented to facilitate production of West Nile virus envelope (WNV-E) fusion proteins in *E. coli*. One approach is to engineer a smaller WNV-E antigen by fusing domain III (EIII) and, optionally, with amino acids of domain II of the WNV-E protein to full-length STF2 (e.g., STF2.E, STF2.EIII+). Domain III is responsible for virus-host interactions and retains many West Nile virus neutralizing antibody epitopes. It also contains only 2 of the 12 cysteine residues present within the full length envelope protein, making expression in *E. coli* more feasible. A second approach has been to delete the hyper-variable hinge region of flagellin (e.g., STF2Δ) thereby creating a smaller fusion protein (STF2Δ.EIII+) The hyper-variable region of flagellin is not required for TLR5 signaling and its removal may also reduce the immunogenic potential of flagellin. Both STF2.EIII+ and STF2Δ.EIII+ have been expressed in *E. coli* and purified. The purified proteins have been characterized for TLR5 signaling activity in bioassays and for E epitope display in ELISA assays using a panel of WNV-E polyclonal and neutralizing monoclonal antibodies. Results from these studies indicate that STF2Δ.EIII+ has higher PAMP activity and more conformation-sensitive neutralizing WNV-E epitopes than STF2.EIII+.

Purity of STF2.EIII+ and STF2Δ.EIII+

Several lots of STF2.EIII+ and STF2Δ.EIII+ have been produced in *E. coli* and purified (Table 12). STF2.EIII+ was expressed as a soluble protein and purified under non-denaturing conditions using a 4-step process, as described above, that included anion exchange chromatography and gel filtration. Final yields from 6 L cultures ranged from about 0.9 mg to about 3.8 mg and all preparations contained low levels of endotoxin as measured by standard LAL procedures (about < 0.1 EU/μg protein, see *supra*). In contrast, STF2Δ.EIII+ formed inclusion bodies in *E. coli*, and was purified under denaturing conditions. All chromatography steps used to purify STF2Δ.EIII+ required the use of 8M urea. Following purification, the denatured protein was refolded by step-wise dialysis to allow for gradual urea removal. Refolding was typically carried out at protein concentrations of about 0.3 mg/ml without any loss due to protein precipitation. Two preparations of STF2Δ.EIII+ from a single 6L culture yielded about 1.2 and about 6.7 mg of protein, both of which had acceptable endotoxin levels. As expected, purified STF2.EIII+ and STF2Δ.EIII+ migrated on SDS PAGE under reducing conditions as about 65 kDa and about 43 kDa proteins, respectively. Notably, STF2Δ.EIII+ migrated slightly faster under non-reducing conditions. This altered migration may be due to disulfide bond formation involving the two cysteines residues in domain III of the envelope protein. As well, a larger species of STF2Δ.EIII+ was detected by Western blot analysis whose molecular weight is consistent with a trimer form of the protein ("(STF2Δ.EIII+)x3 or 3 units of STF2Δ.EIII+").

Table 11: Endotoxin levels and TLR-5 activity for STF2.EIII+ (SEQ ID NO: 658) and STF2Δ.EIII+ (SEQ ID NO: 674) fusion proteins.

Batch Number	Protein	Yield (mg)	Endotoxin Levels (EU/μg)	TLR-5 EC_{50}
052	STF2.EIII+	3.8	0.03	>5000.00 ng/ml
054	STF2.EIII+	0.9	0.02	1195.00 ng/ml
057	STF2.EIII+	1.6	0.07	197.92 ng/ml
044	STF2Δ.EIII+	1.2	0.07	1.13 ng/ml
045	STF2Δ.EIII+	6.7	0.07	4.34 ng/ml

TLR5 activity in the HEK293 IL-8 assay

To compare the PAMP activity of both fusion proteins, a TLR5 bioassay was performed. HEK293 IL-8 cells were treated with serial dilutions of two independent protein batches (Figures 62A and 62B). Cultures were incubated for a 24 hour period and conditioned media were harvested and assayed for IL-8 production by ELISA. As shown in Figure 62A, STF2Δ.EIII+ showed potent TLR-5 activity. Regression analysis of the titration curve determined the EC_{50} of batches 2004-044 and 2004-045 to be 1.13 ng/ml and 4.34 ng/ml, respectively (Table 11, *supra*). In both cases, the TLR5 specific-activity was at least about 10-fold higher than the control protein STF2.OVA. In contrast, 2 preparations of STF2.EIII+ showed significantly weaker TLR5 activity than STF2.OVA. The EC_{50} of STF2.EIII+ batches 054 and 057 were about 1195.00 ng/ml and about 197.92 ng/ml.

Antigenicity of STF2.EIII+ and STF2Δ.EIII+

The antigenicity of STF2.EIII+ and STF2Δ.EIII+ was examined by direct ELISA using a flagellin monoclonal antibody specific for the N-terminal region of STF2 (6H11, Inotek Pharmaceuticals, Beverly, MA) and a panel of WNV-E-specific antibodies (5C5, 5H10, 3A3, 7H2 and 3D9, Bioreliance, Road Rockville, MD) previously shown to neutralize West Nile virus *in vitro*. As shown in Figure 63, a comparison of the reactivity of full length West Nile virus envelope protein with STF2Δ.EIII+ revealed that West Nile virus monoclonal antibodies 5C5, 5H10, 3A3 and 7H2, but not 3D9 recognize the fusion protein. This pattern of reactivity is consistent with the proposed location of 5C5, 5H10, 3A3 and 7H2 epitopes within EIII. The epitope for 3D9 lies outside of domain III of the West Nile virus envelope

protein. As expected, all West Nile virus monoclonal antibodies reacted with full length West Nile virus envelope protein and the flagellin monoclonal only reacted with STF2Δ.EIII+. Both proteins reacted with a polyclonal West Nile virus envelope antiserum, but STF2Δ.EIII+ reactivity was somewhat reduced, perhaps due to the reduced number of potential epitopes present in the smaller domain.

Using 5C5 and 7H10 WNV monoclonal antibodies, a direct antigenic comparison was made between STF2.EIII+ and STF2Δ.EIII+ (Figures 64A, 64B, 64C and 64D). In these studies, plates were coated with the indicated proteins and then detected with polyclonal rabbit anti-E, or mouse monoclonal antibodies as described. As shown in Figures 64A, 64B, 64C and 64D, both STF2.EIII+ and STF2Δ.EIII+ were readily detected with the flagellin monoclonal antibody with no significant differences in reactivity. However, distinct reactivity with the anti-envelope monoclonal antibodies was observed. The reactivity of STF2Δ.EIII+ with either 5C5 or 7H2 was significantly greater than that observed with STF2.EIII+. Collectively, these results indicate that the flagellin 6H11 epitope of STF2Δ.EIII+ is uncompromised and is comparable to the flagellin sequence of STF2.EIII+. They also highlight distinct differences in the antigenicity of the EIII domains of these proteins and indicate that STF2Δ.EIII+ contains more of the critical conformation dependent neutralizing epitopes than STF2.EIII+.

Efficacy and immunogenicity

Several efficacy studies designed to examine the protective efficacy our candidates in C3H/HeN mice following challenge with West Nile virus have been completed. Studies typically consisted of 5 groups of mice (10 mice per group) immunized intraperitoneally (i.p.) or subcutaneously (s.c.) on days 0, 14 and 28. On days 21 and 35, sera were harvested and tested for West Nile virus envelope protein - IgG antibody (ELISA) and the ability to neutralize West Nile virus *in vitro* (PRNT assay). On day 35, mice were challenged with a lethal dose of West Nile virus strain 2741. Survival was monitored for 21 days post-challenge.

Mice were immunized with PBS, *Drosophila* conditioned medium containing STF2.E (CM, positive control), 25 μg of STF2Δ.EIII+ i.p., 25 μg STF2Δ.EIII+ s.c., 25 μg STF2.EIII+ i.p. and 25 μg STF2.EIII+ s.c. The West Nile

virus envelope protein antibody responses and survival data are shown Figures 65 and 66. By day 35 all groups that received STF2Δ.EIII+ had significant levels of West Nile virus envelope protein IgG. In contrast, mice that received STF2.EIII+ had no measurable West Nile virus envelope protein antibody response. Administration of STF2Δ.EIII+ i.p. or s.c led to 100% survival following West Nile virus challenge. Consistent with the poor immunogenicity of STF2.EIII+, little to no protection was provided by this candidate when compared to the PBS control. The poor immunogenicity and efficacy of STF2.EIII+ in this study are attributed to the reduced TLR5 activity and/or the weak EIII epitope reactivity of this protein.

Plaque reduction neutralization titers

To further evaluate the West Nile virus envelope protein antibody response elicited by STF2Δ.EIII+ and potentially correlate protective efficacy with neutralizing antibody titers, the plaque reduction neutralization test (PRNT) was performed. Day 35 serum samples from efficacy studies described above were tested for their ability to block West Nile virus infection in cultured Vero cells. Briefly, pooled mouse serum samples were heat-inactivated and serially diluted two-fold in PBS with 0.5% gelatin. Dilutions starting with 1:10 were incubated with about 100 pfu of the West Nile virus strain 2741. The virus/serum mixture was incubated at about 37°C for 1 h and then inoculated onto confluent monolayers of Vero cells (ATCC, Catalog Number CCL-81, Manassas, Virginia) in duplicate wells of 6-well tissue culture plates. The virus was allowed to adsorb to the cell monolayer prior to adding a 1% agarose overlay. Infected cell cultures were incubated for 4 days at 37°C followed by a second agarose overlay containing 4% neutral red. Virus plaques were counted 12 h later. Serum titers that led to 80% reduction in viral plaque numbers ($PRNT_{80}$) were recorded.

A summary of the $PRNT_{80}$ data from efficacy studies concerning STF2.EIII+ and STF2Δ.EIII+ is presented in Table 12 below. In two independent studies involving STF2.EIII+ where survival of about 50% or less was reported, pooled sera failed to inhibit plaque formation. This finding is not surprising given the weak antibody response elicited by this construct. In three efficacy studies involving STF2Δ.EIII+ where survival was about 70% or greater, pooled sera had

neutralization titers of 1:40 or better. Neutralization titers of 1:40 or greater typically correlate with protection *in vivo*.

Table 12: Survivial and $PRNT_{80}$ Results for STF2.EIII+ (SEQ ID NO: 658), STF2Δ.EIII+ (SEQ ID NO: 674) and STF2.E (SEQ ID NO: 762) CM (Control Media) Fusion Proteins

Batch	Candidate	Route	Study #	Survival (%)	$PRNT_{80}$ (dilution)
054	STF2.EIII+	i.p.	3	50	Negative
057	STF2.EIII+	i.p.	4	11	Negative
057	STF2.EIII+	s.c.	4	20	negative
044	STF2Δ.EIII+	i.p.	2	70	1:40
045	STF2Δ.EIII+	i.p.	3	90	1:40
045	STF2Δ.EIII+	s.c.	3	100	1:160
045	STF2Δ.EIII+	i.p.	4	100	1:80
045	STF2Δ.EIII+	s.c.	4	100	1:40
--	STF2.E CM	i.p.	3	90	1:640
--	STF2.E CM	i.p.	4	--	1:1280

STF2Δ.EIIIs+ a modified version of STF2Δ.EIII+

Protein preparations of STF2Δ.EIII+ tested in the mouse efficacy studies described above were purified by anion-exchange and size-exclusion chromatography steps carried out under denaturing conditions followed by refolding using step-wise dialysis. With this process, two predominant species that correspond to the monomeric and trimeric forms of STF2Δ.EIII+ were generated and present as a mixture in the final product. To minimize the heterogeneity of the final product, new refolding and purification methods have been developed that favor the production of either monomer or trimer. Because it is unclear which form of STF2Δ.EIII+ is the active component or if both are equally potent, both species have been produced in milligram quantities and tested for efficacy in mice.

It was initially unclear as to why STF2Δ.EIII+ refolding resulted in the formation of a trimeric species. However, when the sequence of the STF2Δ.EIII+ expression construct was re-examined, we identified a cysteine residue within the linker sequence that separates STF2Δ from EIII+. The presence of this cysteine would likely interfere with the formation of the appropriate disulfide bond during refolding and might account for the trimeric form of STF2Δ.EIII+. This

unnecessary cysteine was changed to a serine using site-directed mutagenesis and the modified protein (STF2Δ.EIIIs+) was produced and purified. It should be noted that refolding the serine-substituted construct yielded only monomeric protein.

Protective efficacy of STF2Δ.EIII+ (monomer) and STF2Δ.EIIIs+ (trimer) were evaluated in C3H/HeN mice following challenge with West Nile virus. Five groups of mice (10 per group) were immunized with about 25 ug of protein s.c. on days 0, 14 and 28. On days 21 and 35, sera were harvested and tested for WNV-E IgG antibody (ELISA). On day 38, mice were challenged with a lethal dose of WNV strain 2741 and survival was monitored for 21 days. ELISA results from boost 2 (day 35, Figure 67) and survival data (Figure 68) indicate that all constructs elicited significant levels of WNV-E reactive IgG prior to viral challenge and provided about 90% to about 100% protection against the lethal infection. These findings indicate that monomeric or multimeric (e.g., trimers) forms of STFΔ.EIII+ are efficacious and removal of the additional cysteine from the construct does not appreciably impact potency. Removal of the cysteine within the linker sequence may simplify purification of the protein by reducing heterogeneity following protein refolding.

Conclusion

Two recombinant fusion proteins containing the *Salmonella typhimurium* flagellin (STF2) fused to EIII+ domain of West Nile virus envelope protein have been generated. One includes the full length STF2 sequence (STF2.EIII+) and the other a modified version of STF2 that lacks the internal hypervariable region of STF2 (STF2Δ.EIII+). Both proteins have been expressed in *E. coli* and purified by conventional means using anion exchange chromatography and gel filtration. STF2.EIII+ was produced as a soluble protein and was purified under non-denaturing conditions. In contrast, STF2Δ.EIII+ was expressed as an insoluble protein and was purified under denaturing conditions and refolded by step-wise dialysis to remove urea. In HEK293 IL8 assays, preparations of STF2Δ.EIII+ showed greater TLR-5 activity than STF2.EIII+.

In envelope protein epitope display analysis using ELISA assays and West Nile virus envelope prtoein antibodies, STF2Δ.EIII+ displayed more of the critical

conformation dependent neutralizing epitopes. Consistent with the potent TLR-5 activity and envelope protein epitope antigenicity observed with STF2Δ.EIII+, STF2Δ.EIII+ was highly immunogenic and efficacious in mice challenged with a lethal dose of West Nile virus. Because monomeric and trimeric species of STF2Δ.EIII+ were generated during the purification process of this protein, a cysteine within the linker sequence of the expression construct was changed to a serine. Removal of this cysteine eliminated the production of trimeric forms of the protein during refolding and resulted in the generation of monomeric product that displayed potent efficacy *in vivo*.

Japanese encephalitis fusion protein

JE virus is localized in Asia and northern Australia (about 50,000 cases with about 10,000 deaths annually). An approved inactivated virus vaccine was recently associated with a case of acute disseminated encephalomyelitis, prompting the Japanese Ministry of Health, Labor and Welfare to recommend the nationwide suspension of the vaccine. Given the complexities of producing inactivated viruses in infected mouse brains or even in cell culture, and the potential for adverse events associated with inactivated viruses, the opportunity for recombinant-based JE vaccine is appealing.

A STF2Δ.JEIII+ fusion construct was constructed. The JE EIII+ DNA fragment was generated synthetically and codon optimized for expression in *E. coli*. The sequence was ligated into pET24STF2Δ to generate pETSTF2Δ.JEIII+. Expression constructs have been screened by restriction analysis and for expression in *E. coli* BLR(DE3) by IPTG induction. The DNA sequence of each construct has been confirmed, and production of the protein has been scaled up. A batch of material has been generated. A total of about 24 mg of material was purified. This material has potent TLR5 activity, acceptable levels of endotoxin (about 0.03 EU/μg) and a A280/A260 ratio of about 1.3.

Flavivirus peptides

Identification of WNV-E specific antibody epitopes

To identify linear epitopes within the West Nile virus envelope protein that are recognized by antisera from STFΔ.EIIIs+ immunized mice, several synthetic

peptide arrays were generated. One array consisted of overlapping peptides of 20 amino acids in length that spanned the entire West Nile virus domain III and parts of domain II (SEQ ID NOS: 728-754). ELISA results with this array identified a highly reactive 20 amino acid sequence that mapped to the N-terminal region of domain III and included part of the domain I domain CRVKMEKLQLKGTTYGVCSK (SEQ ID NO: 728). To fine map this epitope, additional arrays were generated that focused on the domain I and II junctions (SEQ ID NOS: 703-754). These arrays included an alanine substitution scan to identify amino acids critical for antibody binding (SEQ ID NOS: 728-754). As shown in Figures 69 and 70, antisera from STF2Δ.EIII (monomer and trimer) and STF2Δ.EIIIs+ immunized mice reacted with peptides that spanned the EI/EIII junction (peptides E-30 to E-42) and included the E2-21 peptide CRVKMEKLQLKGTTYGVCSK (SEQ ID NO: 728). This reactivity was severely reduced when specific amino acids (E6, K7, L10 and K11) were changed to alanines (Figure 71). Although it is not known if the antibodies that recognize this epitope are neutralizing, efforts are underway to design and test a peptide vaccine based on this region of WNV-E.

Immunogenicity of Pam3Cys.WNV001 peptide vaccine

A lipidated West Nile virus envelope protein fused to Pam3Cys on the N-terminal end was synthesized using the 20 amino acid sequence LTSGHLKCRVKMEKLQLKGT (SEQ ID NO:772) (Putnak, R., *et al, Vaccine 23*:4442-4452 (2005)). The immunogenicity of this peptide was tested in C3H/HeN mice and compared to peptide without Pam3Cys (Figure 72). The reactivity of antisera from immunized animals was tested by direct ELISA as described in the legend and the results indicate that the Pam3Cys.WNV001 peptide is significantly more immunogenic than the peptide without the TLR2 modification. The antisera from these studies will be tested in virus neutralization assays (PRNT) to determine if the antibodies elicited will neutralize West Nile virus *in vitro*. The lipidated peptide will also be tested in the West Nile virus challenge model to assess protective efficacy against a lethal virus challenge.

Assay Development
Competition ELISA assay development

To assess the neutralizing potential of antisera derived from immunized mice, a competition ELISA assay was developed using well-characterized monoclonal antibody (7H2) that neutralizes West Nile virus in culture and reacts with a conformation-sensitive epitope within the EIII domain of the West Nile virus envelope protein antigen. The assay was designed as a capture ELISA that measures the ability of sera from immunized animals to prevent 7H2 from binding West Nile virus envelope protein. Serial dilutions ranging from 1:10 to 1:5000 of day 35 mouse antisera from efficacy study 4 (Figures 65 and 66, Table 13) were incubated with biotinylated West Nile virus envelope protein and then added to ELISA plates pre-coated with 7H2 monoclonal antibody (Bioreliance, Road Rockville, MD). Following several washes to remove unbound material, bound West Nile virus envelope protein was detected using avidin-HRP. Results from a representative experiment are shown in Figure 69. At dilutions of 1:25, a measurable loss of West Nile virus envelope protein binding to 7H2-coated plates was observed when antisera derived from animals immunized with STF2Δ.EIIIs where tested. No competition was detected with antisera derived from mock immunized animals that received PBS in place of antigen. These initial results demonstrate that antibodies elicited by STF2Δ.EIII+ compete with 7H2 for binding Wests Nile virus envelope protein. These findings are consistent with the protection from WNV infection observed in animals immunized with STF2Δ.EIII+ and help establish a correlation between antibody epitope reactivity *in vitro* and efficacy *in vivo*.

While this invention has been particularly shown and described with references to example embodiments thereof, it will be understood by those skilled in the art that various changes in form and details may be made therein without departing from the scope of the invention encompassed by the appended claims.

What is claimed is:

1. A method of making a protein that stimulates a protective immune response in a subject, comprising the steps of:

a) separating a portion of a protein from a naturally occurring viral hemagglutinin to thereby form a protein portion, wherein the protein portion includes,

i) at least a portion of a globular head, and

ii) at least a portion of at least one secondary structure having at least one β-sheet at a bottom of the globular head that causes the globular head to essentially retain its tertiary structure,

and wherein the protein portion lacks a membrane fusion domain, a transmembrane domain and a cytoplasmic domain;

b) transforming a nucleic acid sequence encoding the protein portion into a prokaryotic host cell;

c) culturing the prokaryotic host cell to thereby make the protein portion; and

d) fusing the protein portion to at least a portion of a flagellin to thereby make the protein that stimulates a protective immune response in the subject.

2. The method of Claim 1, wherein:

i) the protein portion further lacks a signal sequence;

ii) the secondary structure includes at least a portion of at least one ß-pleated sheet or at least one alpha helix;

iii) the protein portion further includes at least one member selected from the group consisting of a salt bridge, a leucine zipper and a zinc finger; and/or

iv) the protein portion further includes at least one disulfide bond, at least two disulfide bonds, at least four disulfide bonds, at least five disulfide bonds or at least six disulfide bonds.

3. The method of Claim 1, further including the step of substituting a nucleic acid sequence encoding at least one amino acid residue selected from the group consisting of a hydrophilic amino acid residue, a polar amino acid residue and a neutral amino acid residue, for a nucleic acid sequence that encodes at least one hydrophobic amino acid residue in the protein portion, wherein the hydrophobic amino acid residue includes at least one member selected from the group consisting of a phenylalanine residue, a tryptophan residue and a tyrosine residue or wherein the polar amino acid residue includes at least one member selected from the group consisting of an aspartic acid residue and a glutamic acid residue.

4. The method of Claim 1, wherein the viral hemagglutinin includes an influenza A, influenza B or influenza C viral hemagglutinin protein and wherein the influenza A is at least one member selected from the group consisting of H1, H2, H3, H5, H7 and H9.

5. The method of Claim 1, wherein the prokaryotic host cell includes an *E. coli* prokaryotic host cell, a *Pseudomonas* prokaryotic host cell, a *Bacillus* prokaryotic host cell, or a *Salmonella* prokaryotic host cell.

6. The method of Claim 1, further including the step of transforming the prokaryotic host cell with a chaperone nucleic acid sequence, wherein the chaperone nucleic acid sequence is at least one member selected from the group consisting of a groES-groEL chaperone, a dnaK-dnaJ-grpE chaperone, a groES-groEL-tig chaperone or a tig chaperone.

7. The method of Claim 1, further including the step of operably linking a nucleic acid sequence encoding the flagellin to the nucleic acid sequence encoding the protein portion of the viral hemagglutinin to thereby fuse the protein portion to the flagellin.

8. The method of Claim 1, further including the step of inserting a linker between the Toll-like Receptor agonist and the protein portion.

9. The method of Claim 7, wherein the flagellin that lacks at least a portion of a hinge region or, wherein the flagellin includes at least one member selected from the group consisting of *Salmonella typhimurium* flagellin, an *E. coli* flagellin, a *S. muenchen* flagellin, a *Yersinia* flagellin, a *P. aeruginosa* flagellin and a *L. monocytogenes* flagellin.

10. The method of Claim 1, further including the step of operably linking a nucleic acid sequence encoding a carrier protein to the nucleic acid sequence encoding the protein portion, wherein the carrier protein is at least one member selected from the group consisting of tetanus toxoid, a *Vibrio cholerae* toxoid, a diphtheria toxoid, a cross-reactive mutant of diphtheria toxoid, a *E. coli* B subunit of a heat labile enterotoxin, a tobacco mosaic virus coat protein, a rabies virus envelope protein, a rabies virus envelope glycoprotein, a thyroglobulin, a heat shock protein 60, a keyhole limpet hemocyanin and an early secreted antigen tuberculosis-6.

11. The method of Claim 1, further including the step of fusing at least a portion of a carrier protein to the protein portion.

12. A method of making a protein that stimulates a protective immune response in a subject, comprising the steps of:

a) separating a portion of a protein from a naturally occurring influenza viral hemagglutinin to thereby form a protein portion, wherein the protein portion includes,

i) at least a portion of a globular head, and

ii) at least a portion of at least one secondary structure having at least one β-sheet at a bottom of the globular head that causes the globular head to essentially retain its tertiary structure,

and wherein the protein portion lacks a membrane fusion domain, a transmembrane domain and a cytoplasmic domain;

b) transfecting a nucleic acid sequence encoding the protein portion into eukaryotic host cell, wherein the eukaryotic host cell is not a *Pichia pastoris* eukaryotic host cell;

c) culturing the eukaryotic host cell to thereby make the protein portion; and

d) fusing the protein portion to at least a portion of a flagellin to thereby make the protein that stimulates a protective immune response in the subject.

13. The method of Claim 12, wherein

i) the protein portion further lacks a signal sequence, further includes a sialic acid binding site; and/or

ii) the secondary structure includes at least a portion of at least one ß-pleated sheet, wherein the secondary structure includes at least one member selected from the group consisting of a salt bridge and an alpha helix.

14. The method of Claim 12, further including substituting a nucleic acid sequence encoding at least one amino acid residue selected from the group consisting of a hydrophilic amino acid residue, a polar amino acid residue and a neutral amino acid residue, for a nucleic acid sequence that encodes at least one hydrophobic amino acid residue in the protein portion.

15. The method of Claim 12, further including the step of deleting at least one glycosylation site in at least one of the nucleic acid sequence encoding the protein portion or the flagellin, wherein the glycosylation site includes an N-glycosylation site.

16. The method of Claim 12, wherein the eukaryotic host cell includes a *Saccharomyces* eukaryotic host cell, a fungal eukaryotic host cell, a *Kluyveromyces lactis* host cell or a parasite eukaryotic host cell, wherein the parasite eukaryotic host cell includes a *Leishmania tarentolae* eukaryotic host cell.

17. The method of Claim 12, further including the step of operably linking a nucleic acid sequence encoding the flagellin to the nucleic acid sequence encoding the protein portion to thereby fuse the protein portion to the flagellin.

18. The method of Claim 12, further including the step of inserting a linker between the flagellin and the protein portion.

19. The method of Claim 12, wherein the flagellin that lacks at least a portion of a hinge region, or wherein the flagellin includes at least one member selected from the group consisting of *Salmonella typhimurium* flagellin, a *E. coli* flagellin, a *S. muenchen* flagellin, a *Yersinia* flagellin, a *P. aerugtnosa* flagellin and a *L. monocytogenes* flagellin.

20. The method of Claim 19, further including the step of deleting at least one glycosylation site in the nucleic acid sequence encoding the flagellin.

21. Use of a composition that includes at least a portion of a flagellin fused to a protein portion of a naturally occurring viral hemagglutinin, wherein the protein portion includes at least a portion of a globular head and at least a portion of one secondary structure having at least one β-sheet at a bottom of the globular head that causes the globular head to essentially retain its tertiary structure, and lacks a membrane fusion domain, a transmembrane domain and a cytoplasmic domain for the manufacture of a medicament for stimulating protective immunity in a subject.

22. A method of making a viral hemagglutinin protein that stimulates a protective immune response in a subject, comprising the steps of:

a) separating a portion of a protein from a naturally occurring influenza viral hemagglutinin to thereby form a protein portion, wherein the protein portion includes,

i) at least a portion of a globular head, and

ii) at least a portion of at least one secondary structure having at least one β-sheet at a bottom of the globular head that causes the globular head to essentially retain its tertiary structure, and wherein the protein portion

lacks a membrane fusion domain, transmembrane domain and a cytoplasmic domain;

b) transfecting a nucleic acid sequence encoding the protein portion in a eukaryotic host cell, wherein the eukaryotic host cell is neither a *Pichia pastoris* eukaryotic host cell nor a stably transfected insect host cell;

c) culturing the eukaryotic host cell to thereby make the protein portion; and

d) fusing the protein portion to at least a portion of a flagellin to thereby make the protein that stimulates a protective immune response in the subject.

23. A method of making a viral hemagglutinin protein that stimulates a protective immune response in a subject, comprising the steps of:

a) transforming a prokaryotic host cell with a nucleic acid sequence encoding at least one influenza viral hemagglutinin that includes,

i) at least a portion of a globular head, and

ii) at least a portion of at least one secondary structure having at least one β-sheet at a bottom of the globular head that causes the globular head to essentially retain its tertiary structure, wherein the viral hemagglutinin lacks a membrane fusion domain, a transmembrane domain and a cytoplasmic domain;

b) culturing the prokaryotic cell to thereby make the protein portion; and

c) fusing the protein portion to at least a portion of a flagellin to thereby make the protein that stimulates a protective immune response in the subject.

24. A composition comprising at least a portion of a flagellin and a portion of a protein of a naturally occurring viral hemagglutinin, wherein the portion of the protein of the naturally occurring viral hemagglutinin includes at least a portion of a globular head and at least a portion of at least one secondary structure having at least one β-sheet at a bottom of the globular head that causes the globular head to essentially retain its tertiary structure, and wherein the portion of the naturally

occurring viral hemagglutinin lacks a membrane fusion domain, a transmembrane domain, and a cytoplasmic domain.

25. The composition of Claim 24, further including a carrier protein.

26. The composition of Claim 24, wherein the flagellin and the portion of the naturally occurring viral hemagglutinin are components of a fusion protein.

27. The composition of Claim 24, wherein the flagellin lacks at least a portion of a hinge region or, wherein the Toll-like Receptor 5 agonist includes a flagellin and includes at least one member selected from the group consisting of *Salmonella* typhimurium flagellin, an *E coli* flagellin, a *S. muenchen* flagellin, a *Yersinia* flagellin, a *P. aerugenosa* flagellin and a *L. monocytogenes* flagellin.

28. The method of any of Claims 1-20, 22 and 23, wherein the portion of the protein of the naturally occurring influenza viral hemagglutinin further includes at least one β-sandwich at the bottom of the portion of the globular head.

29. The method of Claim 28, wherein the portion of the protein of the naturally occurring influenza viral hemagglutinin further includes at least two β-strands at the bottom of the portion of the globular head.

30. The method of any of Claims 1-20, 22, 23, 28 and 29, wherein:

i) the flagellin and the portion of the naturally occurring influenza viral hemagglutinin are components of a fusion protein that includes an amino acid linker between the flagellin and the portion of the naturally occurring influenza viral hemagglutinin; and

ii) the naturally occurring influenza viral hemagglutinin is a H3 strain.

31. The method of any of Claims 1-20, 22, 23, 28, 29, and 30, wherein the portion of the naturally occurring influenza viral hemagglutinin lacks an H2 subunit.

32. The composition of any of Claims 24-27, wherein the portion of the protein of the naturally occurring influenza viral hemagglutinin further includes at least one β-sandwich at the bottom of the portion of the globular head.

33. The composition of Claim 32, wherein the portion of the protein of the naturally occurring influenza viral hemagglutinin further includes at least two β-strands at the bottom of the portion of the globular head.

34. The composition of any of Claims 24-27, 32 and 33, wherein:

i) the flagellin and the portion of the naturally occurring influenza viral hemagglutinin are components of a fusion protein that includes an amino acid linker between the flagellin and the portion of the naturally occurring influenza viral hemagglutinin; and

ii) the naturally occurring influenza viral hemagglutinin is a H3 strain.

35. The composition of any of Claims 24-27 and 32-34, wherein the portion of the naturally occurring influenza viral hemagglutinin lacks an H2 subunit.

36. The protein made by the method of any of claims 1-20, 22, 23, 28, 29, 30 or 31 for use in stimulating protective immunity in a subject to a viral hemagglutinin infection.

37. Use of the protein made by the method of any of claims 1-20, 22, 23, 28, 29, 30 or 31 in the manufacture of a medicament for stimulating protective immunity towards a viral hemagglutinin infection.

38. The method of any of claims 1-20, 22, 23, 28, 29, 30 or 31, wherein the portion of the naturally occurring viral hemagglutinin includes an HA1-1 portion of an influenza viral hemagglutinin and/or an HA1-2 portion of an influenza viral hemagglutinin.

39. The composition of any of claims 24-27, 32, 33, 34 or 35, wherein the portion of the naturally occurring viral hemagglutinin includes an HA1-1 portion of an

influenza viral hemagglutinin and/or an HA1-2 portion of an influenza viral hemagglutinin.

SEQUENCE LISTING

<110> Song, Langzhou
Nakaar, Valerian
Price, Albert E.
Tussey, Lynda G.
Huleatt, James W.
Powell, Thomas J.
Evans, Robert K.

<120> COMPOSITIONS THAT INCLUDE HEMAGGLUTININ, METHODS OF MAKING AND METHODS OF USE THEREOF

<130> 3710.1021002

<150> 60/779,854
<151> 2006-03-07

<150> 60/784,497
<151> 2006-03-20

<150> 60/790,457
<151> 2006-04-07

<150> 60/814,292
<151> 2006-06-16

<150> 60/830,881
<151> 2006-07-14

<150> 60/838,007
<151> 2006-08-16

<150> 60/856,451
<151> 2006-11-03

<160> 841

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 565
<212> PRT
<213> Influenza A/Puerto Rico/8/34 HA

<400> 1

```
Met Lys Ala Asn Leu Leu Val Leu Leu Ser Ala Leu Ala Ala Ala Asp
 1               5                  10                  15
Ala Asp Thr Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Asp Thr
            20                  25                  30
Val Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His Ser Val Asn
        35                  40                  45
Leu Leu Glu Asp Ser His Asn Gly Lys Leu Cys Arg Leu Lys Gly Ile
    50                  55                  60
Ala Pro Leu Gln Leu Gly Lys Cys Asn Ile Ala Gly Trp Leu Leu Gly
65                  70                  75                  80
Asn Pro Glu Cys Asp Pro Leu Leu Pro Val Arg Ser Trp Ser Tyr Ile
```

```
                85                  90                  95
Val Glu Thr Pro Asn Ser Glu Asn Gly Ile Cys Tyr Pro Gly Asp Phe
            100                 105                 110
Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe
        115                 120                 125
Glu Arg Phe Glu Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His Asn
    130                 135                 140
Thr Asn Gly Val Thr Ala Ala Cys Ser His Glu Gly Lys Ser Ser Phe
145                 150                 155                 160
Tyr Arg Asn Leu Leu Trp Leu Thr Glu Lys Glu Gly Ser Tyr Pro Lys
                165                 170                 175
Leu Lys Asn Ser Tyr Val Asn Lys Lys Gly Lys Glu Val Leu Val Leu
            180                 185                 190
Trp Gly Ile His His Pro Pro Asn Ser Lys Glu Gln Gln Asn Leu Tyr
        195                 200                 205
Gln Asn Glu Asn Ala Tyr Val Ser Val Val Thr Ser Asn Tyr Asn Arg
    210                 215                 220
Arg Phe Thr Pro Glu Ile Ala Glu Arg Pro Lys Val Arg Asp Gln Ala
225                 230                 235                 240
Gly Arg Met Asn Tyr Tyr Trp Thr Leu Leu Lys Pro Gly Asp Thr Ile
                245                 250                 255
Ile Phe Glu Ala Asn Gly Asn Leu Ile Ala Pro Met Tyr Ala Phe Ala
            260                 265                 270
Leu Ser Arg Gly Phe Gly Ser Gly Ile Ile Thr Ser Asn Ala Ser Met
        275                 280                 285
His Glu Cys Asn Thr Lys Cys Gln Thr Pro Leu Gly Ala Ile Asn Ser
    290                 295                 300
Ser Leu Pro Tyr Gln Asn Ile His Pro Val Thr Ile Gly Glu Cys Pro
305                 310                 315                 320
Lys Tyr Val Arg Ser Ala Lys Leu Arg Met Val Thr Gly Leu Arg Asn
                325                 330                 335
Ile Pro Ser Ile Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe
            340                 345                 350
Ile Glu Gly Gly Trp Thr Gly Met Ile Asp Gly Trp Tyr Gly Tyr His
        355                 360                 365
His Gln Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Gln Lys Ser Thr
    370                 375                 380
Gln Asn Ala Ile Asn Gly Ile Thr Asn Lys Val Asn Thr Val Ile Glu
385                 390                 395                 400
Lys Met Asn Ile Gln Phe Thr Ala Val Gly Lys Glu Phe Asn Lys Leu
                405                 410                 415
Glu Lys Arg Met Glu Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Leu
            420                 425                 430
Asp Ile Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu
        435                 440                 445
Arg Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys
    450                 455                 460
Val Lys Ser Gln Leu Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys
465                 470                 475                 480
Phe Glu Phe Tyr His Lys Cys Asp Asn Glu Cys Met Glu Ser Val Arg
                485                 490                 495
Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ser Lys Leu Asn
            500                 505                 510
Arg Glu Lys Val Asp Gly Val Lys Leu Glu Ser Met Gly Ile Tyr Gln
        515                 520                 525
Ile Leu Ala Ile Tyr Ser Thr Val Ala Ser Ser Leu Val Leu Leu Val
    530                 535                 540
Ser Leu Gly Ala Ile Ser Phe Trp Met Cys Ser Asn Gly Ser Leu Gln
545                 550                 555                 560
```

Cys Arg Ile Cys Ile
565

<210> 2
<211> 564
<212> PRT
<213> Influenza A/Vietnam/1203/2004 HA

<400> 2
Met Glu Lys Ile Val Leu Leu Phe Ala Ile Val Ser Leu Val Lys Ser
1 5 10 15
Asp Gln Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Glu Gln Val
20 25 30
Asp Thr Ile Met Glu Lys Asn Val Thr Val Thr His Ala Gln Asp Ile
35 40 45
Leu Glu Lys Lys His Asn Gly Lys Leu Cys Asp Leu Asp Gly Val Lys
50 55 60
Pro Leu Ile Leu Arg Asp Cys Ser Val Ala Gly Trp Leu Leu Gly Asn
65 70 75 80
Pro Met Cys Asp Glu Phe Ile Asn Val Pro Glu Trp Ser Tyr Ile Val
85 90 95
Glu Lys Ala Asn Pro Val Asn Asp Leu Cys Tyr Pro Gly Asp Phe Asn
100 105 110
Asp Tyr Glu Glu Leu Lys His Leu Leu Ser Arg Ile Asn His Phe Glu
115 120 125
Lys Ile Gln Ile Ile Pro Lys Ser Ser Trp Ser Ser His Glu Ala Ser
130 135 140
Leu Gly Val Ser Ser Ala Cys Pro Tyr Gln Gly Lys Ser Ser Phe Phe
145 150 155 160
Arg Asn Val Val Trp Leu Ile Lys Lys Asn Ser Thr Tyr Pro Thr Ile
165 170 175
Lys Arg Ser Tyr Asn Asn Thr Asn Gln Glu Asp Leu Leu Val Leu Trp
180 185 190
Gly Ile His His Pro Asn Asp Ala Ala Glu Gln Thr Lys Leu Tyr Gln
195 200 205
Asn Pro Thr Thr Tyr Ile Ser Val Gly Thr Ser Thr Leu Asn Gln Arg
210 215 220
Leu Val Pro Arg Ile Ala Thr Arg Ser Lys Val Asn Gly Gln Ser Gly
225 230 235 240
Arg Met Glu Phe Phe Trp Thr Ile Leu Lys Pro Asn Asp Ala Ile Asn
245 250 255
Phe Glu Ser Asn Gly Asn Phe Ile Ala Pro Glu Tyr Ala Tyr Lys Ile
260 265 270
Val Lys Lys Gly Asp Ser Thr Ile Met Lys Ser Glu Leu Glu Tyr Gly
275 280 285
Asn Cys Asn Thr Lys Cys Gln Thr Pro Met Gly Ala Ile Asn Ser Ser
290 295 300
Met Pro Phe His Asn Ile His Pro Leu Thr Ile Gly Glu Cys Pro Lys
305 310 315 320
Tyr Val Lys Ser Asn Arg Leu Val Leu Ala Thr Gly Leu Arg Asn Ser
325 330 335
Pro Gln Arg Glu Arg Arg Arg Lys Lys Arg Gly Leu Phe Gly Ala Ile
340 345 350
Ala Gly Phe Ile Glu Gly Gly Trp Gln Gly Met Val Asp Gly Trp Tyr
355 360 365
Gly Tyr His His Ser Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Lys
370 375 380
Glu Ser Thr Gln Lys Ala Ile Asp Gly Val Thr Asn Lys Val Asn Ser

```
385                 390                 395                 400
Ile Ile Asp Lys Met Asn Thr Gln Phe Glu Ala Val Gly Arg Glu Phe
                405                 410                 415
Asn Asn Leu Glu Arg Arg Ile Glu Asn Leu Asn Lys Lys Met Glu Asp
            420                 425                 430
Gly Phe Leu Asp Val Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Met
        435                 440                 445
Glu Asn Glu Arg Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu
    450                 455                 460
Tyr Asp Lys Val Arg Leu Gln Leu Arg Asp Asn Ala Lys Glu Leu Gly
465                 470                 475                 480
Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp Asn Glu Cys Met Glu
                485                 490                 495
Ser Val Arg Asn Gly Thr Asp Tyr Pro Gln Tyr Ser Glu Glu Ala Arg
            500                 505                 510
Leu Lys Arg Glu Glu Ile Ser Gly Val Lys Leu Glu Ser Ile Gly Ile
        515                 520                 525
Tyr Gln Ile Leu Ser Ile Tyr Ser Thr Val Ala Ser Ser Leu Ala Leu
    530                 535                 540
Ala Ile Met Val Ala Gly Leu Ser Leu Trp Met Cys Ser Asn Gly Ser
545                 550                 555                 560
Leu Gln Cys Arg
```

<210> 3
<211> 568
<212> PRT
<213> Influenza A/Indonesia/5/2005 HA

<400> 3

```
Met Glu Lys Ile Val Leu Leu Leu Ala Ile Val Ser Leu Val Lys Ser
 1                5                  10                  15
Asp Gln Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Glu Gln Val
            20                  25                  30
Asp Thr Ile Met Glu Lys Asn Val Thr Val Thr His Ala Gln Asp Ile
        35                  40                  45
Leu Glu Lys Thr His Asn Gly Lys Leu Cys Asp Leu Asp Gly Val Lys
    50                  55                  60
Pro Leu Ile Leu Arg Asp Cys Ser Val Ala Gly Trp Leu Leu Gly Asn
65                  70                  75                  80
Pro Met Cys Asp Glu Phe Ile Asn Val Pro Glu Trp Ser Tyr Ile Val
                85                  90                  95
Glu Lys Ala Asn Pro Thr Asn Asp Leu Cys Tyr Pro Gly Ser Phe Asn
            100                 105                 110
Asp Tyr Glu Glu Leu Lys His Leu Leu Ser Arg Ile Asn His Phe Glu
        115                 120                 125
Lys Ile Gln Ile Ile Pro Lys Ser Ser Trp Ser Asp His Glu Ala Ser
    130                 135                 140
Ser Gly Val Ser Ser Ala Cys Pro Tyr Leu Gly Ser Pro Ser Phe Phe
145                 150                 155                 160
Arg Asn Val Val Trp Leu Ile Lys Lys Asn Ser Thr Tyr Pro Thr Ile
                165                 170                 175
Lys Lys Ser Tyr Asn Asn Thr Asn Gln Glu Asp Leu Leu Val Leu Trp
            180                 185                 190
Gly Ile His His Pro Asn Asp Ala Ala Glu Gln Thr Arg Leu Tyr Gln
        195                 200                 205
Asn Pro Thr Thr Tyr Ile Ser Ile Gly Thr Ser Thr Leu Asn Gln Arg
    210                 215                 220
```

```
Leu Val Pro Lys Ile Ala Thr Arg Ser Lys Val Asn Gly Gln Ser Gly
225                 230                 235                 240
Arg Met Glu Phe Phe Trp Thr Ile Leu Lys Pro Asn Asp Ala Ile Asn
                245                 250                 255
Phe Glu Ser Asn Gly Asn Phe Ile Ala Pro Glu Tyr Ala Tyr Lys Ile
            260                 265                 270
Val Lys Lys Gly Asp Ser Ala Ile Met Lys Ser Glu Leu Glu Tyr Gly
        275                 280                 285
Asn Cys Asn Thr Lys Cys Gln Thr Pro Met Gly Ala Ile Asn Ser Ser
    290                 295                 300
Met Pro Phe His Asn Ile His Pro Leu Thr Ile Gly Glu Cys Pro Lys
305                 310                 315                 320
Tyr Val Lys Ser Asn Arg Leu Val Leu Ala Thr Gly Leu Arg Asn Ser
                325                 330                 335
Pro Gln Arg Glu Ser Arg Arg Lys Lys Arg Gly Leu Phe Gly Ala Ile
            340                 345                 350
Ala Gly Phe Ile Glu Gly Gly Trp Gln Gly Met Val Asp Gly Trp Tyr
        355                 360                 365
Gly Tyr His His Ser Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Lys
    370                 375                 380
Glu Ser Thr Gln Lys Ala Ile Asp Gly Val Thr Asn Lys Val Asn Ser
385                 390                 395                 400
Ile Ile Asp Lys Met Asn Thr Gln Phe Glu Ala Val Gly Arg Glu Phe
                405                 410                 415
Asn Asn Leu Glu Arg Arg Ile Glu Asn Leu Asn Lys Lys Met Glu Asp
            420                 425                 430
Gly Phe Leu Asp Val Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Met
        435                 440                 445
Glu Asn Glu Arg Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu
    450                 455                 460
Tyr Asp Lys Val Arg Leu Gln Leu Arg Asp Asn Ala Lys Glu Leu Gly
465                 470                 475                 480
Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp Asn Glu Cys Met Glu
                485                 490                 495
Ser Ile Arg Asn Gly Thr Tyr Asn Tyr Pro Gln Tyr Ser Glu Glu Ala
            500                 505                 510
Arg Leu Lys Arg Glu Glu Ile Ser Gly Val Lys Leu Glu Ser Ile Gly
        515                 520                 525
Thr Tyr Gln Ile Leu Ser Ile Tyr Ser Thr Val Ala Ser Ser Leu Ala
    530                 535                 540
Leu Ala Ile Met Met Ala Gly Leu Ser Leu Trp Met Cys Ser Asn Gly
545                 550                 555                 560
Ser Leu Gln Cys Arg Ile Cys Ile
                565
```

<210> 4
<211> 565
<212> PRT
<213> Influenza A/NewCaledonia/20/1999

<400> 4

```
Met Lys Ala Lys Leu Leu Val Leu Leu Cys Thr Phe Thr Ala Thr Tyr
 1               5                  10                  15
Ala Asp Thr Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Asp Thr
            20                  25                  30
Val Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His Ser Val Asn
        35                  40                  45
Leu Leu Glu Asp Ser His Asn Gly Lys Leu Cys Leu Leu Lys Gly Ile
```

50 55 60
Ala Pro Leu Gln Leu Gly Asn Cys Ser Val Ala Gly Trp Ile Leu Gly
65 70 75 80
Asn Pro Glu Cys Glu Leu Leu Ile Ser Lys Glu Ser Trp Ser Tyr Ile
85 90 95
Val Glu Thr Pro Asn Pro Glu Asn Gly Thr Cys Tyr Pro Gly Tyr Phe
100 105 110
Ala Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe
115 120 125
Glu Arg Phe Glu Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His Thr
130 135 140
Val Thr Gly Val Ser Ala Ser Cys Ser His Asn Gly Lys Ser Ser Phe
145 150 155 160
Tyr Arg Asn Leu Leu Trp Leu Thr Gly Lys Asn Gly Leu Tyr Pro Asn
165 170 175
Leu Ser Lys Ser Tyr Val Asn Asn Lys Glu Lys Glu Val Leu Val Leu
180 185 190
Trp Gly Val His His Pro Pro Asn Ile Gly Asn Gln Arg Ala Leu Tyr
195 200 205
His Thr Glu Asn Ala Tyr Val Ser Val Val Ser Ser His Tyr Ser Arg
210 215 220
Arg Phe Thr Pro Glu Ile Ala Lys Arg Pro Lys Val Arg Asp Gln Glu
225 230 235 240
Gly Arg Ile Asn Tyr Tyr Trp Thr Leu Leu Glu Pro Gly Asp Thr Ile
245 250 255
Ile Phe Glu Ala Asn Gly Asn Leu Ile Ala Pro Trp Tyr Ala Phe Ala
260 265 270
Leu Ser Arg Gly Phe Gly Ser Gly Ile Ile Thr Ser Asn Ala Pro Met
275 280 285
Asp Glu Cys Asp Ala Lys Cys Gln Thr Pro Gln Gly Ala Ile Asn Ser
290 295 300
Ser Leu Pro Phe Gln Asn Val His Pro Val Thr Ile Gly Glu Cys Pro
305 310 315 320
Lys Tyr Val Arg Ser Ala Lys Leu Arg Met Val Thr Gly Leu Arg Asn
325 330 335
Ile Pro Ser Ile Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe
340 345 350
Ile Glu Gly Gly Trp Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr His
355 360 365
His Gln Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Gln Lys Ser Thr
370 375 380
Gln Asn Ala Ile Asn Gly Ile Thr Asn Lys Val Asn Ser Val Ile Glu
385 390 395 400
Lys Met Asn Thr Gln Phe Thr Ala Val Gly Lys Glu Phe Asn Lys Leu
405 410 415
Glu Arg Arg Met Glu Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Leu
420 425 430
Asp Ile Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu
435 440 445
Arg Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys
450 455 460
Val Lys Ser Gln Leu Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys
465 470 475 480
Phe Glu Phe Tyr His Lys Cys Asn Asn Glu Cys Met Glu Ser Val Lys
485 490 495
Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ser Lys Leu Asn
500 505 510
Arg Glu Lys Ile Asp Gly Val Lys Leu Glu Ser Met Gly Val Tyr Gln
515 520 525

```
Ile Leu Ala Ile Tyr Ser Thr Val Ala Ser Ser Leu Val Leu Leu Val
    530                 535                 540
Ser Leu Gly Ala Ile Ser Phe Trp Met Cys Ser Asn Gly Ser Leu Gln
545                 550                 555                 560
Cys Arg Ile Cys Ile
                565


<210> 5
<211> 566
<212> PRT
<213> Influenza A/South Carolina/1/18HA

<400> 5
Met Glu Ala Arg Leu Leu Val Leu Leu Cys Ala Phe Ala Ala Thr Asn
 1               5                  10                  15
Ala Asp Thr Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Asp Thr
            20                  25                  30
Val Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His Ser Val Asn
        35                  40                  45
Leu Leu Glu Asp Ser His Asn Gly Lys Leu Cys Lys Leu Lys Gly Ile
    50                  55                  60
Ala Pro Leu Gln Leu Gly Lys Cys Asn Ile Ala Gly Trp Leu Leu Gly
65                  70                  75                  80
Asn Pro Glu Cys Asp Leu Leu Leu Thr Ala Ser Ser Trp Ser Tyr Ile
                85                  90                  95
Val Glu Thr Ser Asn Ser Glu Asn Gly Thr Cys Tyr Pro Gly Asp Phe
            100                 105                 110
Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe
        115                 120                 125
Glu Lys Phe Glu Ile Phe Pro Lys Thr Ser Ser Trp Pro Asn His Glu
    130                 135                 140
Thr Thr Lys Gly Val Thr Ala Ala Cys Ser Tyr Ala Gly Ala Ser Ser
145                 150                 155                 160
Phe Tyr Arg Asn Leu Leu Trp Leu Thr Lys Lys Gly Ser Ser Tyr Pro
                165                 170                 175
Lys Leu Ser Lys Ser Tyr Val Asn Asn Lys Gly Lys Glu Val Leu Val
            180                 185                 190
Leu Trp Gly Val His His Pro Pro Thr Gly Thr Asp Gln Gln Ser Leu
        195                 200                 205
Tyr Gln Asn Ala Asp Ala Tyr Val Ser Val Gly Ser Ser Lys Tyr Asn
    210                 215                 220
Arg Arg Phe Thr Pro Glu Ile Ala Ala Arg Pro Lys Val Arg Asp Gln
225                 230                 235                 240
Ala Gly Arg Met Asn Tyr Tyr Trp Thr Leu Leu Glu Pro Gly Asp Thr
                245                 250                 255
Ile Thr Phe Glu Ala Thr Gly Asn Leu Ile Ala Pro Trp Tyr Ala Phe
            260                 265                 270
Ala Leu Asn Arg Gly Ser Gly Ser Gly Ile Ile Thr Ser Asp Ala Pro
        275                 280                 285
Val His Asp Cys Asn Thr Lys Cys Gln Thr Pro His Gly Ala Ile Asn
    290                 295                 300
Ser Ser Leu Pro Phe Gln Asn Ile His Pro Val Thr Ile Gly Glu Cys
305                 310                 315                 320
Pro Lys Tyr Val Arg Ser Thr Lys Leu Arg Met Ala Thr Gly Leu Arg
                325                 330                 335
Asn Ile Pro Ser Ile Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly
            340                 345                 350
Phe Ile Glu Gly Gly Trp Thr Gly Met Ile Asp Gly Trp Tyr Gly Tyr
```

```
        355                 360                 365
His His Gln Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Gln Lys Ser
    370                 375                 380
Thr Gln Asn Ala Ile Asp Gly Ile Thr Asn Lys Val Asn Ser Val Ile
385                 390                 395                 400
Glu Lys Met Asn Thr Gln Phe Thr Ala Val Gly Lys Glu Phe Asn Asn
                405                 410                 415
Leu Glu Arg Arg Ile Glu Asn Leu Asn Lys Lys Val Asp Asp Gly Phe
            420                 425                 430
Leu Asp Ile Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Leu Glu Asn
        435                 440                 445
Glu Arg Thr Leu Asp Phe His Asp Ser Asn Val Arg Asn Leu Tyr Glu
    450                 455                 460
Lys Val Lys Ser Gln Leu Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly
465                 470                 475                 480
Cys Phe Glu Phe Tyr His Lys Cys Asp Asp Ala Cys Met Glu Ser Val
                485                 490                 495
Arg Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ser Lys Leu
            500                 505                 510
Asn Arg Glu Glu Ile Asp Gly Val Lys Leu Glu Ser Met Gly Val Tyr
        515                 520                 525
Gln Ile Leu Ala Ile Tyr Ser Thr Val Ala Ser Ser Leu Val Leu Leu
    530                 535                 540
Val Ser Leu Gly Ala Ile Ser Phe Trp Met Cys Ser Asn Gly Ser Leu
545                 550                 555                 560
Gln Cys Arg Ile Cys Ile
                565
```

<210> 6
<211> 329
<212> PRT
<213> Influenza A/Wisconsin/67/2005 HA

<400> 6

```
Gln Lys Leu Pro Gly Asn Asp Asn Ser Thr Ala Thr Leu Cys Leu Gly
 1                5                  10                  15
His His Ala Val Pro Asn Gly Thr Ile Val Lys Thr Ile Thr Asn Asp
            20                  25                  30
Gln Ile Glu Val Thr Asn Ala Thr Glu Leu Val Gln Ser Ser Ser Thr
        35                  40                  45
Gly Gly Ile Cys Asp Ser Pro His Gln Ile Leu Asp Gly Glu Asn Cys
    50                  55                  60
Thr Leu Ile Asp Ala Leu Leu Gly Asp Pro Gln Cys Asp Gly Phe Gln
65                  70                  75                  80
Asn Lys Lys Trp Asp Leu Phe Val Glu Arg Ser Lys Ala Tyr Ser Asn
                85                  90                  95
Cys Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Leu Arg Ser Leu Val
            100                 105                 110
Ala Ser Ser Gly Thr Leu Glu Phe Asn Asp Glu Ser Phe Asn Trp Thr
        115                 120                 125
Gly Val Thr Gln Asn Gly Thr Ser Ser Ala Cys Lys Arg Arg Ser Asn
    130                 135                 140
Asn Ser Phe Phe Ser Arg Leu Asn Trp Leu Thr His Leu Lys Phe Lys
145                 150                 155                 160
Tyr Pro Ala Leu Asn Val Thr Met Pro Asn Asn Glu Lys Phe Asp Lys
                165                 170                 175
Leu Tyr Ile Trp Gly Val His His Pro Gly Thr Asp Asn Asp Gln Ile
            180                 185                 190
```

Phe Leu His Ala Gln Ala Ser Gly Arg Ile Thr Val Ser Thr Lys Arg
195 200 205
Ser Gln Gln Thr Val Ile Pro Asn Ile Gly Ser Arg Pro Arg Ile Arg
210 215 220
Asn Ile Pro Ser Arg Ile Ser Ile Tyr Trp Thr Ile Val Lys Pro Gly
225 230 235 240
Asp Ile Leu Leu Ile Asn Ser Thr Gly Asn Leu Ile Ala Pro Arg Gly
245 250 255
Tyr Phe Lys Ile Arg Ser Gly Lys Ser Ser Ile Met Arg Ser Asp Ala
260 265 270
Pro Ile Gly Lys Cys Asn Ser Glu Cys Ile Thr Pro Asn Gly Ser Ile
275 280 285
Pro Asn Asp Lys Pro Phe Gln Asn Val Asn Arg Ile Thr Tyr Gly Ala
290 295 300
Cys Pro Arg Tyr Val Lys Gln Asn Thr Leu Lys Leu Ala Thr Gly Met
305 310 315 320
Arg Asn Val Pro Glu Lys Gln Thr Arg
325

<210> 7
<211> 328
<212> PRT
<213> Influenza A/X31 subtype H3N2 HA

<400> 7
Gln Asp Leu Pro Gly Asn Asp Asn Ser Thr Ala Thr Leu Cys Leu Gly
1 5 10 15
His His Ala Val Pro Asn Gly Thr Leu Val Lys Thr Ile Thr Asp Asp
20 25 30
Gln Ile Glu Val Thr Asn Ala Thr Glu Leu Val Gln Ser Ser Ser Thr
35 40 45
Gly Lys Ile Cys Asn Asn Pro His Arg Ile Leu Asp Gly Ile Asp Cys
50 55 60
Thr Leu Ile Asp Ala Leu Leu Gly Asp Pro His Cys Asp Val Phe Gln
65 70 75 80
Asn Glu Thr Trp Asp Leu Phe Val Glu Arg Ser Lys Ala Phe Ser Asn
85 90 95
Cys Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Leu Arg Ser Leu Val
100 105 110
Ala Ser Ser Gly Thr Leu Glu Phe Ile Thr Glu Gly Phe Thr Trp Thr
115 120 125
Gly Val Ile Gln Asn Gly Gly Ser Asn Ala Cys Lys Arg Gly Pro Gly
130 135 140
Ser Gly Phe Phe Ser Arg Leu Asn Trp Leu Thr Lys Ser Gly Ser Thr
145 150 155 160
Tyr Pro Val Leu Asn Val Thr Met Pro Asn Asn Asp Asn Phe Asp Lys
165 170 175
Leu Tyr Ile Trp Gly Ile His His Pro Ser Thr Asn Gln Glu Gln Thr
180 185 190
Ser Leu Tyr Val Gln Ala Ser Gly Arg Val Thr Val Ser Thr Arg Arg
195 200 205
Ser Gln Gln Thr Ile Ile Pro Asn Ile Gly Ser Arg Pro Trp Val Arg
210 215 220
Gly Leu Ser Ser Arg Ile Ser Ile Tyr Trp Thr Ile Val Lys Pro Gly
225 230 235 240
Asp Val Leu Val Ile Asn Ser Asn Gly Asn Leu Ile Ala Pro Arg Gly
245 250 255
Tyr Phe Lys Met Arg Thr Gly Lys Ser Ser Ile Met Arg Ser Asp Ala

260 265 270
Pro Ile Asp Thr Cys Ile Ser Glu Cys Ile Thr Pro Asn Gly Ser Ile
275 280 285
Pro Asn Asp Lys Pro Phe Gln Asn Val Asn Lys Ile Thr Tyr Gly Ala
290 295 300
Cys Pro Lys Tyr Val Lys Gln Asn Thr Leu Lys Leu Ala Thr Gly Met
305 310 315 320
Arg Asn Val Pro Glu Lys Gln Thr
325

<210> 8
<211> 272
<212> PRT
<213> Artificial Sequence

<220>
<223> PR/8 HA1-1

<400> 8
Ser His Asn Gly Lys Leu Cys Arg Leu Lys Gly Ile Ala Pro Leu Gln
1 5 10 15
Leu Gly Lys Cys Asn Ile Ala Gly Trp Leu Leu Gly Asn Pro Glu Cys
20 25 30
Asp Pro Leu Leu Pro Val Arg Ser Trp Ser Tyr Ile Val Glu Thr Pro
35 40 45
Asn Ser Glu Asn Gly Ile Cys Tyr Pro Gly Asp Phe Ile Asp Tyr Glu
50 55 60
Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe Glu Arg Phe Glu
65 70 75 80
Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His Asn Thr Asn Gly Val
85 90 95
Thr Ala Ala Cys Ser His Glu Gly Lys Ser Ser Phe Tyr Arg Asn Leu
100 105 110
Leu Trp Leu Thr Glu Lys Glu Gly Ser Tyr Pro Lys Leu Lys Asn Ser
115 120 125
Tyr Val Asn Lys Lys Gly Lys Glu Val Leu Val Leu Trp Gly Ile His
130 135 140
His Pro Pro Asn Ser Lys Glu Gln Gln Asn Leu Tyr Gln Asn Glu Asn
145 150 155 160
Ala Tyr Val Ser Val Val Thr Ser Asn Tyr Asn Arg Arg Phe Thr Pro
165 170 175
Glu Ile Ala Glu Arg Pro Lys Val Arg Asp Gln Ala Gly Arg Met Asn
180 185 190
Tyr Tyr Trp Thr Leu Leu Lys Pro Gly Asp Thr Ile Ile Phe Glu Ala
195 200 205
Asn Gly Asn Leu Ile Ala Pro Met Tyr Ala Phe Ala Leu Ser Arg Gly
210 215 220
Phe Gly Ser Gly Ile Ile Thr Ser Asn Ala Ser Met His Glu Cys Asn
225 230 235 240
Thr Lys Cys Gln Thr Pro Leu Gly Ala Ile Asn Ser Ser Leu Pro Tyr
245 250 255
Gln Asn Ile His Pro Val Thr Ile Gly Glu Cys Pro Lys Tyr Val Arg
260 265 270

<210> 9
<211> 223
<212> PRT

<213> Artificial Sequence

<220>
<223> PR/8 HA1-2

<400> 9
Lys Gly Ile Ala Pro Leu Gln Leu Gly Lys Cys Asn Ile Ala Gly Trp
1 5 10 15
Leu Leu Gly Asn Pro Glu Cys Asp Pro Leu Leu Pro Val Arg Ser Trp
20 25 30
Ser Tyr Ile Val Glu Thr Pro Asn Ser Glu Asn Gly Ile Cys Tyr Pro
35 40 45
Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val
50 55 60
Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Glu Ser Ser Trp Pro
65 70 75 80
Asn His Asn Thr Asn Gly Val Thr Ala Ala Cys Ser His Glu Gly Lys
85 90 95
Ser Ser Phe Tyr Arg Asn Leu Leu Trp Leu Thr Glu Lys Glu Gly Ser
100 105 110
Tyr Pro Lys Leu Lys Asn Ser Tyr Val Asn Lys Lys Gly Lys Glu Val
115 120 125
Leu Val Leu Trp Gly Ile His His Pro Pro Asn Ser Lys Glu Gln Gln
130 135 140
Asn Leu Tyr Gln Asn Glu Asn Ala Tyr Val Ser Val Val Thr Ser Asn
145 150 155 160
Tyr Asn Arg Arg Phe Thr Pro Glu Ile Ala Glu Arg Pro Lys Val Arg
165 170 175
Asp Gln Ala Gly Arg Met Asn Tyr Tyr Trp Thr Leu Leu Lys Pro Gly
180 185 190
Asp Thr Ile Ile Phe Glu Ala Asn Gly Asn Leu Ile Ala Pro Met Tyr
195 200 205
Ala Phe Ala Leu Ser Arg Gly Phe Gly Ser Gly Ile Ile Thr Ser
210 215 220

<210> 10
<211> 176
<212> PRT
<213> Artificial Sequence

<220>
<223> PR/8 HA1-3

<400> 10
Asn Ser Glu Asn Gly Ile Cys Tyr Pro Gly Asp Phe Ile Asp Tyr Glu
1 5 10 15
Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe Glu Arg Phe Glu
20 25 30
Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His Asn Thr Asn Gly Val
35 40 45
Thr Ala Ala Cys Ser His Glu Gly Lys Ser Ser Phe Tyr Arg Asn Leu
50 55 60
Leu Trp Leu Thr Glu Lys Glu Gly Ser Tyr Pro Lys Leu Lys Asn Ser
65 70 75 80
Tyr Val Asn Lys Lys Gly Lys Glu Val Leu Val Leu Trp Gly Ile His
85 90 95
His Pro Pro Asn Ser Lys Glu Gln Gln Asn Leu Tyr Gln Asn Glu Asn
100 105 110

```
Ala Tyr Val Ser Val Val Thr Ser Asn Tyr Asn Arg Arg Phe Thr Pro
         115                 120                 125
Glu Ile Ala Glu Arg Pro Lys Val Arg Asp Gln Ala Gly Arg Met Asn
     130                 135                 140
Tyr Tyr Trp Thr Leu Leu Lys Pro Gly Asp Thr Ile Ile Phe Glu Ala
145                 150                 155                 160
Asn Gly Asn Leu Ile Ala Pro Met Tyr Ala Phe Ala Leu Ser Arg Gly
                165                 170                 175
```

<210> 11
<211> 274
<212> PRT
<213> Artificial Sequence

<220>
<223> VN HA1-1

<400> 11

```
Glu Lys Lys His Asn Gly Lys Leu Cys Asp Leu Asp Gly Val Lys Pro
 1               5                  10                  15
Leu Ile Leu Arg Asp Cys Ser Val Ala Gly Trp Leu Leu Gly Asn Pro
            20                  25                  30
Met Cys Asp Glu Phe Ile Asn Val Pro Glu Trp Ser Tyr Ile Val Glu
         35                  40                  45
Lys Ala Asn Pro Val Asn Asp Leu Cys Tyr Pro Gly Asp Phe Asn Asp
     50                  55                  60
Tyr Glu Glu Leu Lys His Leu Leu Ser Arg Ile Asn His Phe Glu Lys
65                  70                  75                  80
Ile Gln Ile Ile Pro Lys Ser Ser Trp Ser Ser His Glu Ala Ser Leu
                85                  90                  95
Gly Val Ser Ser Ala Cys Pro Tyr Gln Gly Lys Ser Ser Phe Phe Arg
            100                 105                 110
Asn Val Val Trp Leu Ile Lys Lys Asn Ser Thr Tyr Pro Thr Ile Lys
         115                 120                 125
Arg Ser Tyr Asn Asn Thr Asn Gln Glu Asp Leu Leu Val Leu Trp Gly
     130                 135                 140
Ile His His Pro Asn Asp Ala Ala Glu Gln Thr Lys Leu Tyr Gln Asn
145                 150                 155                 160
Pro Thr Thr Tyr Ile Ser Val Gly Thr Ser Thr Leu Asn Gln Arg Leu
                165                 170                 175
Val Pro Arg Ile Ala Thr Arg Ser Lys Val Asn Gly Gln Ser Gly Arg
            180                 185                 190
Met Glu Phe Phe Trp Thr Ile Leu Lys Pro Asn Asp Ala Ile Asn Phe
         195                 200                 205
Glu Ser Asn Gly Asn Phe Ile Ala Pro Glu Tyr Ala Tyr Lys Ile Val
     210                 215                 220
Lys Lys Gly Asp Ser Thr Ile Met Lys Ser Glu Leu Glu Tyr Gly Asn
225                 230                 235                 240
Cys Asn Thr Lys Cys Gln Thr Pro Met Gly Ala Ile Asn Ser Ser Met
                245                 250                 255
Pro Phe His Asn Ile His Pro Leu Thr Ile Gly Glu Cys Pro Lys Tyr
            260                 265                 270
Val Lys
```

<210> 12
<211> 223

<212> PRT
<213> Artificial Sequence

<220>
<223> VN HA1-2

<400> 12
Gly Val Lys Pro Leu Ile Leu Arg Asp Cys Ser Val Ala Gly Trp Leu
1 5 10 15
Leu Gly Asn Pro Met Cys Asp Glu Phe Ile Asn Val Pro Glu Trp Ser
20 25 30
Tyr Ile Val Glu Lys Ala Asn Pro Val Asn Asp Leu Cys Tyr Pro Gly
35 40 45
Asp Phe Asn Asp Tyr Glu Glu Leu Lys His Leu Leu Ser Arg Ile Asn
50 55 60
His Phe Glu Lys Ile Gln Ile Ile Pro Lys Ser Ser Trp Ser Ser His
65 70 75 80
Glu Ala Ser Leu Gly Val Ser Ser Ala Cys Pro Tyr Gln Gly Lys Ser
85 90 95
Ser Phe Phe Arg Asn Val Val Trp Leu Ile Lys Lys Asn Ser Thr Tyr
100 105 110
Pro Thr Ile Lys Arg Ser Tyr Asn Asn Thr Asn Gln Glu Asp Leu Leu
115 120 125
Val Leu Trp Gly Ile His His Pro Asn Asp Ala Ala Glu Gln Thr Lys
130 135 140
Leu Tyr Gln Asn Pro Thr Thr Tyr Ile Ser Val Gly Thr Ser Thr Leu
145 150 155 160
Asn Gln Arg Leu Val Pro Arg Ile Ala Thr Arg Ser Lys Val Asn Gly
165 170 175
Gln Ser Gly Arg Met Glu Phe Phe Trp Thr Ile Leu Lys Pro Asn Asp
180 185 190
Ala Ile Asn Phe Glu Ser Asn Gly Asn Phe Ile Ala Pro Glu Tyr Ala
195 200 205
Tyr Lys Ile Val Lys Lys Gly Asp Ser Thr Ile Met Lys Ser Glu
210 215 220

<210> 13
<211> 174
<212> PRT
<213> Artificial Sequence

<220>
<223> VN HA1-3

<400> 13
Asn Asp Leu Cys Tyr Pro Gly Asp Phe Asn Asp Tyr Glu Glu Leu Lys
1 5 10 15
His Leu Leu Ser Arg Ile Asn His Phe Glu Lys Ile Gln Ile Ile Pro
20 25 30
Lys Ser Ser Trp Ser Ser His Glu Ala Ser Leu Gly Val Ser Ser Ala
35 40 45
Cys Pro Tyr Gln Gly Lys Ser Ser Phe Phe Arg Asn Val Val Trp Leu
50 55 60
Ile Lys Lys Asn Ser Thr Tyr Pro Thr Ile Lys Arg Ser Tyr Asn Asn
65 70 75 80
Thr Asn Gln Glu Asp Leu Leu Val Leu Trp Gly Ile His His Pro Asn
85 90 95
Asp Ala Ala Glu Gln Thr Lys Leu Tyr Gln Asn Pro Thr Thr Tyr Ile

```
            100                 105                 110
Ser Val Gly Thr Ser Thr Leu Asn Gln Arg Leu Val Pro Arg Ile Ala
        115                 120                 125
Thr Arg Ser Lys Val Asn Gly Gln Ser Gly Arg Met Glu Phe Phe Trp
    130                 135                 140
Thr Ile Leu Lys Pro Asn Asp Ala Ile Asn Phe Glu Ser Asn Gly Asn
145                 150                 155                 160
Phe Ile Ala Pro Glu Tyr Ala Tyr Lys Ile Val Lys Lys Gly
                165                 170
```

<210> 14
<211> 274
<212> PRT
<213> Artificial Sequence

<220>
<223> IND HA1-1

<400> 14

```
Glu Lys Thr His Asn Gly Lys Leu Cys Asp Leu Asp Gly Val Lys Pro
 1               5                  10                  15
Leu Ile Leu Arg Asp Cys Ser Val Ala Gly Trp Leu Leu Gly Asn Pro
            20                  25                  30
Met Cys Asp Glu Phe Ile Asn Val Pro Glu Trp Ser Tyr Ile Val Glu
        35                  40                  45
Lys Ala Asn Pro Thr Asn Asp Leu Cys Tyr Pro Gly Ser Phe Asn Asp
    50                  55                  60
Tyr Glu Glu Leu Lys His Leu Leu Ser Arg Ile Asn His Phe Glu Lys
65                  70                  75                  80
Ile Gln Ile Ile Pro Lys Ser Ser Trp Ser Asp His Glu Ala Ser Ser
                85                  90                  95
Gly Val Ser Ser Ala Cys Pro Tyr Leu Gly Ser Pro Ser Phe Phe Arg
            100                 105                 110
Asn Val Val Trp Leu Ile Lys Lys Asn Ser Thr Tyr Pro Thr Ile Lys
        115                 120                 125
Lys Ser Tyr Asn Asn Thr Asn Gln Glu Asp Leu Leu Val Leu Trp Gly
    130                 135                 140
Ile His His Pro Asn Asp Ala Ala Glu Gln Thr Arg Leu Tyr Gln Asn
145                 150                 155                 160
Pro Thr Thr Tyr Ile Ser Ile Gly Thr Ser Thr Leu Asn Gln Arg Leu
                165                 170                 175
Val Pro Lys Ile Ala Thr Arg Ser Lys Val Asn Gly Gln Ser Gly Arg
            180                 185                 190
Met Glu Phe Phe Trp Thr Ile Leu Lys Pro Asn Asp Ala Ile Asn Phe
        195                 200                 205
Glu Ser Asn Gly Asn Phe Ile Ala Pro Glu Tyr Ala Tyr Lys Ile Val
    210                 215                 220
Lys Lys Gly Asp Ser Ala Ile Met Lys Ser Glu Leu Glu Tyr Gly Asn
225                 230                 235                 240
Cys Asn Thr Lys Cys Gln Thr Pro Met Gly Ala Ile Asn Ser Ser Met
                245                 250                 255
Pro Phe His Asn Ile His Pro Leu Thr Ile Gly Glu Cys Pro Lys Tyr
            260                 265                 270
Val Lys
```

<210> 15

<211> 223
<212> PRT
<213> Artificial Sequence

<220>
<223> IND HA1-2

<400> 15

```
Gly Val Lys Pro Leu Ile Leu Arg Asp Cys Ser Val Ala Gly Trp Leu
 1               5                  10                  15
Leu Gly Asn Pro Met Cys Asp Glu Phe Ile Asn Val Pro Glu Trp Ser
            20                  25                  30
Tyr Ile Val Glu Lys Ala Asn Pro Thr Asn Asp Leu Cys Tyr Pro Gly
        35                  40                  45
Ser Phe Asn Asp Tyr Glu Glu Leu Lys His Leu Leu Ser Arg Ile Asn
    50                  55                  60
His Phe Glu Lys Ile Gln Ile Ile Pro Lys Ser Ser Trp Ser Asp His
65                  70                  75                  80
Glu Ala Ser Ser Gly Val Ser Ser Ala Cys Pro Tyr Leu Gly Ser Pro
                85                  90                  95
Ser Phe Phe Arg Asn Val Val Trp Leu Ile Lys Lys Asn Ser Thr Tyr
            100                 105                 110
Pro Thr Ile Lys Lys Ser Tyr Asn Asn Thr Asn Gln Glu Asp Leu Leu
        115                 120                 125
Val Leu Trp Gly Ile His His Pro Asn Asp Ala Ala Glu Gln Thr Arg
    130                 135                 140
Leu Tyr Gln Asn Pro Thr Thr Tyr Ile Ser Ile Gly Thr Ser Thr Leu
145                 150                 155                 160
Asn Gln Arg Leu Val Pro Lys Ile Ala Thr Arg Ser Lys Val Asn Gly
                165                 170                 175
Gln Ser Gly Arg Met Glu Phe Phe Trp Thr Ile Leu Lys Pro Asn Asp
            180                 185                 190
Ala Ile Asn Phe Glu Ser Asn Gly Asn Phe Ile Ala Pro Glu Tyr Ala
        195                 200                 205
Tyr Lys Ile Val Lys Lys Gly Asp Ser Ala Ile Met Lys Ser Glu
    210                 215                 220
```

<210> 16
<211> 174
<212> PRT
<213> Artificial Sequence

<220>
<223> IND HA1-3

<400> 16

```
Asn Asp Leu Cys Tyr Pro Gly Ser Phe Asn Asp Tyr Glu Glu Leu Lys
 1               5                  10                  15
His Leu Leu Ser Arg Ile Asn His Phe Glu Lys Ile Gln Ile Ile Pro
            20                  25                  30
Lys Ser Ser Trp Ser Asp His Glu Ala Ser Ser Gly Val Ser Ser Ala
        35                  40                  45
Cys Pro Tyr Leu Gly Ser Pro Ser Phe Phe Arg Asn Val Val Trp Leu
    50                  55                  60
Ile Lys Lys Asn Ser Thr Tyr Pro Thr Ile Lys Lys Ser Tyr Asn Asn
65                  70                  75                  80
Thr Asn Gln Glu Asp Leu Leu Val Leu Trp Gly Ile His His Pro Asn
                85                  90                  95
```

```
Asp Ala Ala Glu Gln Thr Arg Leu Tyr Gln Asn Pro Thr Thr Tyr Ile
            100                 105                 110
Ser Ile Gly Thr Ser Thr Leu Asn Gln Arg Leu Val Pro Lys Ile Ala
        115                 120                 125
Thr Arg Ser Lys Val Asn Gly Gln Ser Gly Arg Met Glu Phe Phe Trp
    130                 135                 140
Thr Ile Leu Lys Pro Asn Asp Ala Ile Asn Phe Glu Ser Asn Gly Asn
145                 150                 155                 160
Phe Ile Ala Pro Glu Tyr Ala Tyr Lys Ile Val Lys Lys Gly
                165                 170
```

<210> 17
<211> 272
<212> PRT
<213> Artificial Sequence

<220>
<223> NC HA1-1

<400> 17

```
Ser His Asn Gly Lys Leu Cys Leu Leu Lys Gly Ile Ala Pro Leu Gln
 1               5                  10                  15
Leu Gly Asn Cys Ser Val Ala Gly Trp Ile Leu Gly Asn Pro Glu Cys
            20                  25                  30
Glu Leu Leu Ile Ser Lys Glu Ser Trp Ser Tyr Ile Val Glu Thr Pro
        35                  40                  45
Asn Pro Glu Asn Gly Thr Cys Tyr Pro Gly Tyr Phe Ala Asp Tyr Glu
    50                  55                  60
Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe Glu Arg Phe Glu
65                  70                  75                  80
Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His Thr Val Thr Gly Val
                85                  90                  95
Ser Ala Ser Cys Ser His Asn Gly Lys Ser Ser Phe Tyr Arg Asn Leu
            100                 105                 110
Leu Trp Leu Thr Gly Lys Asn Gly Leu Tyr Pro Asn Leu Ser Lys Ser
        115                 120                 125
Tyr Val Asn Asn Lys Glu Lys Glu Val Leu Val Leu Trp Gly Val His
    130                 135                 140
His Pro Pro Asn Ile Gly Asn Gln Arg Ala Leu Tyr His Thr Glu Asn
145                 150                 155                 160
Ala Tyr Val Ser Val Val Ser Ser His Tyr Ser Arg Arg Phe Thr Pro
                165                 170                 175
Glu Ile Ala Lys Arg Pro Lys Val Arg Asp Gln Glu Gly Arg Ile Asn
            180                 185                 190
Tyr Tyr Trp Thr Leu Leu Glu Pro Gly Asp Thr Ile Ile Phe Glu Ala
        195                 200                 205
Asn Gly Asn Leu Ile Ala Pro Trp Tyr Ala Phe Ala Leu Ser Arg Gly
    210                 215                 220
Phe Gly Ser Gly Ile Ile Thr Ser Asn Ala Pro Met Asp Glu Cys Asp
225                 230                 235                 240
Ala Lys Cys Gln Thr Pro Gln Gly Ala Ile Asn Ser Ser Leu Pro Phe
                245                 250                 255
Gln Asn Val His Pro Val Thr Ile Gly Glu Cys Pro Lys Tyr Val Arg
            260                 265                 270
```

<210> 18
<211> 223

<212> PRT
<213> Artificial Sequence

<220>
<223> NC HA1-2

<400> 18
Lys Gly Ile Ala Pro Leu Gln Leu Gly Asn Cys Ser Val Ala Gly Trp
1 5 10 15
Ile Leu Gly Asn Pro Glu Cys Glu Leu Leu Ile Ser Lys Glu Ser Trp
20 25 30
Ser Tyr Ile Val Glu Thr Pro Asn Pro Glu Asn Gly Thr Cys Tyr Pro
35 40 45
Gly Tyr Phe Ala Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val
50 55 60
Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Glu Ser Ser Trp Pro
65 70 75 80
Asn His Thr Val Thr Gly Val Ser Ala Ser Cys Ser His Asn Gly Lys
85 90 95
Ser Ser Phe Tyr Arg Asn Leu Leu Trp Leu Thr Gly Lys Asn Gly Leu
100 105 110
Tyr Pro Asn Leu Ser Lys Ser Tyr Val Asn Asn Lys Glu Lys Glu Val
115 120 125
Leu Val Leu Trp Gly Val His His Pro Pro Asn Ile Gly Asn Gln Arg
130 135 140
Ala Leu Tyr His Thr Glu Asn Ala Tyr Val Ser Val Val Ser Ser His
145 150 155 160
Tyr Ser Arg Arg Phe Thr Pro Glu Ile Ala Lys Arg Pro Lys Val Arg
165 170 175
Asp Gln Glu Gly Arg Ile Asn Tyr Tyr Trp Thr Leu Leu Glu Pro Gly
180 185 190
Asp Thr Ile Ile Phe Glu Ala Asn Gly Asn Leu Ile Ala Pro Trp Tyr
195 200 205
Ala Phe Ala Leu Ser Arg Gly Phe Gly Ser Gly Ile Ile Thr Ser
210 215 220

<210> 19
<211> 176
<212> PRT
<213> Artificial Sequence

<220>
<223> NC HA1-3

<400> 19
Asn Pro Glu Asn Gly Thr Cys Tyr Pro Gly Tyr Phe Ala Asp Tyr Glu
1 5 10 15
Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe Glu Arg Phe Glu
20 25 30
Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His Thr Val Thr Gly Val
35 40 45
Ser Ala Ser Cys Ser His Asn Gly Lys Ser Ser Phe Tyr Arg Asn Leu
50 55 60
Leu Trp Leu Thr Gly Lys Asn Gly Leu Tyr Pro Asn Leu Ser Lys Ser
65 70 75 80
Tyr Val Asn Asn Lys Glu Lys Glu Val Leu Val Leu Trp Gly Val His
85 90 95
His Pro Pro Asn Ile Gly Asn Gln Arg Ala Leu Tyr His Thr Glu Asn

```
            100                 105                 110
Ala Tyr Val Ser Val Val Ser Ser His Tyr Ser Arg Arg Phe Thr Pro
        115                 120                 125
Glu Ile Ala Lys Arg Pro Lys Val Arg Asp Gln Glu Gly Arg Ile Asn
    130                 135                 140
Tyr Tyr Trp Thr Leu Leu Glu Pro Gly Asp Thr Ile Ile Phe Glu Ala
145                 150                 155                 160
Asn Gly Asn Leu Ile Ala Pro Trp Tyr Ala Phe Ala Leu Ser Arg Gly
                165                 170                 175


<210> 20
<211> 267
<212> PRT
<213> Artificial Sequence

<220>
<223> WIS HA1-1

<400> 20
Gln Ser Ser Ser Thr Gly Gly Ile Cys Asp Ser Pro His Gln Ile Leu
 1               5                  10                  15
Asp Gly Glu Asn Cys Thr Leu Ile Asp Ala Leu Leu Gly Asp Pro Gln
            20                  25                  30
Cys Asp Gly Phe Gln Asn Lys Lys Trp Asp Leu Phe Val Glu Arg Ser
        35                  40                  45
Lys Ala Tyr Ser Asn Cys Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser
    50                  55                  60
Leu Arg Ser Leu Val Ala Ser Ser Gly Thr Leu Glu Phe Asn Asp Glu
65                  70                  75                  80
Ser Phe Asn Trp Thr Gly Val Thr Gln Asn Gly Thr Ser Ser Ala Cys
                85                  90                  95
Lys Arg Arg Ser Asn Asn Ser Phe Phe Ser Arg Leu Asn Trp Leu Thr
            100                 105                 110
His Leu Lys Phe Lys Tyr Pro Ala Leu Asn Val Thr Met Pro Asn Asn
        115                 120                 125
Glu Lys Phe Asp Lys Leu Tyr Ile Trp Gly Val His His Pro Gly Thr
    130                 135                 140
Asp Asn Asp Gln Ile Phe Leu His Ala Gln Ala Ser Gly Arg Ile Thr
145                 150                 155                 160
Val Ser Thr Lys Arg Ser Gln Gln Thr Val Ile Pro Asn Ile Gly Ser
                165                 170                 175
Arg Pro Arg Ile Arg Asn Ile Pro Ser Arg Ile Ser Ile Tyr Trp Thr
            180                 185                 190
Ile Val Lys Pro Gly Asp Ile Leu Leu Ile Asn Ser Thr Gly Asn Leu
        195                 200                 205
Ile Ala Pro Arg Gly Tyr Phe Lys Ile Arg Ser Gly Lys Ser Ser Ile
    210                 215                 220
Met Arg Ser Asp Ala Pro Ile Gly Lys Cys Asn Ser Glu Cys Ile Thr
225                 230                 235                 240
Pro Asn Gly Ser Ile Pro Asn Asp Lys Pro Phe Gln Asn Val Asn Arg
                245                 250                 255
Ile Thr Tyr Gly Ala Cys Pro Arg Tyr Val Lys
            260                 265


<210> 21
<211> 218
<212> PRT
```

<213> Artificial Sequence

<220>
<223> WIS HA1-2

<400> 21
Ser Pro His Gln Ile Leu Asp Gly Glu Asn Cys Thr Leu Ile Asp Ala
1 5 10 15
Leu Leu Gly Asp Pro Gln Cys Asp Gly Phe Gln Asn Lys Lys Trp Asp
20 25 30
Leu Phe Val Glu Arg Ser Lys Ala Tyr Ser Asn Cys Tyr Pro Tyr Asp
35 40 45
Val Pro Asp Tyr Ala Ser Leu Arg Ser Leu Val Ala Ser Ser Gly Thr
50 55 60
Leu Glu Phe Asn Asp Glu Ser Phe Asn Trp Thr Gly Val Thr Gln Asn
65 70 75 80
Gly Thr Ser Ser Ala Cys Lys Arg Arg Ser Asn Asn Ser Phe Phe Ser
85 90 95
Arg Leu Asn Trp Leu Thr His Leu Lys Phe Lys Tyr Pro Ala Leu Asn
100 105 110
Val Thr Met Pro Asn Asn Glu Lys Phe Asp Lys Leu Tyr Ile Trp Gly
115 120 125
Val His His Pro Gly Thr Asp Asn Asp Gln Ile Phe Leu His Ala Gln
130 135 140
Ala Ser Gly Arg Ile Thr Val Ser Thr Lys Arg Ser Gln Gln Thr Val
145 150 155 160
Ile Pro Asn Ile Gly Ser Arg Pro Arg Ile Arg Asn Ile Pro Ser Arg
165 170 175
Ile Ser Ile Tyr Trp Thr Ile Val Lys Pro Gly Asp Ile Leu Leu Ile
180 185 190
Asn Ser Thr Gly Asn Leu Ile Ala Pro Arg Gly Tyr Phe Lys Ile Arg
195 200 205
Ser Gly Lys Ser Ser Ile Met Arg Ser Asp
210 215

<210> 22
<211> 169
<212> PRT
<213> Artificial Sequence

<220>
<223> WIS HA1-3

<400> 22
Ser Asn Cys Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Leu Arg Ser
1 5 10 15
Leu Val Ala Ser Ser Gly Thr Leu Glu Phe Asn Asp Glu Ser Phe Asn
20 25 30
Trp Thr Gly Val Thr Gln Asn Gly Thr Ser Ser Ala Cys Lys Arg Arg
35 40 45
Ser Asn Asn Ser Phe Phe Ser Arg Leu Asn Trp Leu Thr His Leu Lys
50 55 60
Phe Lys Tyr Pro Ala Leu Asn Val Thr Met Pro Asn Asn Glu Lys Phe
65 70 75 80
Asp Lys Leu Tyr Ile Trp Gly Val His His Pro Gly Thr Asp Asn Asp
85 90 95
Gln Ile Phe Leu His Ala Gln Ala Ser Gly Arg Ile Thr Val Ser Thr
100 105 110

```
Lys Arg Ser Gln Gln Thr Val Ile Pro Asn Ile Gly Ser Arg Pro Arg
        115                 120                 125
Ile Arg Asn Ile Pro Ser Arg Ile Ser Ile Tyr Trp Thr Ile Val Lys
    130                 135                 140
Pro Gly Asp Ile Leu Leu Ile Asn Ser Thr Gly Asn Leu Ile Ala Pro
145                 150                 155                 160
Arg Gly Tyr Phe Lys Ile Arg Ser Gly
                165
```

<210> 23
<211> 511
<212> PRT
<213> Influenza A/Puerto Rico/8/34 HA0s

<400> 23

```
Asp Thr Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Asp Thr Val
 1               5                  10                  15
Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His Ser Val Asn Leu
            20                  25                  30
Leu Glu Asp Ser His Asn Gly Lys Leu Cys Arg Leu Lys Gly Ile Ala
        35                  40                  45
Pro Leu Gln Leu Gly Lys Cys Asn Ile Ala Gly Trp Leu Leu Gly Asn
    50                  55                  60
Pro Glu Cys Asp Pro Leu Leu Pro Val Arg Ser Trp Ser Tyr Ile Val
65                  70                  75                  80
Glu Thr Pro Asn Ser Glu Asn Gly Ile Cys Tyr Pro Gly Asp Phe Ile
                85                  90                  95
Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe Glu
            100                 105                 110
Arg Phe Glu Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His Asn Thr
        115                 120                 125
Asn Gly Val Thr Ala Ala Cys Ser His Glu Gly Lys Ser Ser Phe Tyr
    130                 135                 140
Arg Asn Leu Leu Trp Leu Thr Glu Lys Glu Gly Ser Tyr Pro Lys Leu
145                 150                 155                 160
Lys Asn Ser Tyr Val Asn Lys Lys Gly Lys Glu Val Leu Val Leu Trp
                165                 170                 175
Gly Ile His His Pro Pro Asn Ser Lys Glu Gln Gln Asn Leu Tyr Gln
            180                 185                 190
Asn Glu Asn Ala Tyr Val Ser Val Val Thr Ser Asn Tyr Asn Arg Arg
        195                 200                 205
Phe Thr Pro Glu Ile Ala Glu Arg Pro Lys Val Arg Asp Gln Ala Gly
    210                 215                 220
Arg Met Asn Tyr Tyr Trp Thr Leu Leu Lys Pro Gly Asp Thr Ile Ile
225                 230                 235                 240
Phe Glu Ala Asn Gly Asn Leu Ile Ala Pro Met Tyr Ala Phe Ala Leu
                245                 250                 255
Ser Arg Gly Phe Gly Ser Gly Ile Ile Thr Ser Asn Ala Ser Met His
            260                 265                 270
Glu Cys Asn Thr Lys Cys Gln Thr Pro Leu Gly Ala Ile Asn Ser Ser
        275                 280                 285
Leu Pro Tyr Gln Asn Ile His Pro Val Thr Ile Gly Glu Cys Pro Lys
    290                 295                 300
Tyr Val Arg Ser Ala Lys Leu Arg Met Val Thr Gly Leu Arg Asn Ile
305                 310                 315                 320
Pro Ser Ile Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile
                325                 330                 335
Glu Gly Gly Trp Thr Gly Met Ile Asp Gly Trp Tyr Gly Tyr His His
```

```
            340                 345                 350
Gln Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Gln Lys Ser Thr Gln
        355                 360                 365
Asn Ala Ile Asn Gly Ile Thr Asn Lys Val Asn Thr Val Ile Glu Lys
    370                 375                 380
Met Asn Ile Gln Phe Thr Ala Val Gly Lys Glu Phe Asn Lys Leu Glu
385                 390                 395                 400
Lys Arg Met Glu Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Leu Asp
                405                 410                 415
Ile Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu Arg
            420                 425                 430
Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys Val
        435                 440                 445
Lys Ser Gln Leu Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys Phe
    450                 455                 460
Glu Phe Tyr His Lys Cys Asp Asn Glu Cys Met Glu Ser Val Arg Asn
465                 470                 475                 480
Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ser Lys Leu Asn Arg
                485                 490                 495
Glu Lys Val Asp Gly Val Lys Leu Glu Ser Met Gly Ile Tyr Gln
            500                 505                 510


<210> 24
<211> 223
<212> PRT
<213> Artificial Sequence

<220>
<223> PR/8 HA1-2mut

<400> 24
Lys Gly Ala Ala Pro Leu Gln Leu Gly Lys Cys Asn Ile Ala Gly Trp
 1               5                  10                  15
Leu Leu Gly Asn Pro Glu Cys Asp Pro Leu Leu Pro Val Arg Ser Trp
            20                  25                  30
Ser Asp Ile Ala Glu Thr Pro Asn Ser Glu Asn Gly Ile Cys Tyr Pro
        35                  40                  45
Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val
    50                  55                  60
Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Glu Ser Ser Trp Pro
65                  70                  75                  80
Asn His Asn Thr Asn Gly Val Thr Ala Ala Cys Ser His Glu Gly Lys
                85                  90                  95
Ser Ser Phe Tyr Arg Asn Leu Leu Trp Leu Thr Glu Lys Glu Gly Ser
            100                 105                 110
Tyr Pro Lys Leu Lys Asn Ser Tyr Val Asn Lys Lys Gly Lys Glu Val
        115                 120                 125
Leu Val Leu Trp Gly Ile His His Pro Pro Asn Ser Lys Glu Gln Gln
    130                 135                 140
Asn Leu Tyr Gln Asn Glu Asn Ala Tyr Val Ser Val Val Thr Ser Asn
145                 150                 155                 160
Tyr Asn Arg Arg Phe Thr Pro Glu Ile Ala Glu Arg Pro Lys Val Arg
                165                 170                 175
Asp Gln Ala Gly Arg Met Asn Tyr Tyr Trp Thr Leu Leu Lys Pro Gly
            180                 185                 190
Asp Thr Ile Ile Phe Glu Ala Asn Gly Asn Leu Ile Ala Pro Met Tyr
        195                 200                 205
Ala Phe Ala Leu Ser Arg Gly Phe Gly Ser Gly Ile Ile Thr Ser
```

210 215 220

<210> 25
<211> 176
<212> PRT
<213> Artificial Sequence

<220>
<223> PR/8 HA1-3mut

<400> 25
Asn Ser Glu Asn Glu Ile Cys Tyr Pro Gly Asp Phe Ile Asp Lys Glu
1 5 10 15
Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe Glu Arg Phe Glu
20 25 30
Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His Asn Thr Asn Gly Val
35 40 45
Thr Ala Ala Cys Ser His Glu Gly Lys Ser Ser Phe Tyr Arg Asn Leu
50 55 60
Leu Trp Leu Thr Glu Lys Glu Gly Ser Tyr Pro Lys Leu Lys Asn Ser
65 70 75 80
Tyr Val Asn Lys Lys Gly Lys Glu Val Leu Val Leu Trp Gly Ile His
85 90 95
His Pro Pro Asn Ser Lys Glu Gln Gln Asn Leu Tyr Gln Asn Glu Asn
100 105 110
Ala Tyr Val Ser Val Val Thr Ser Asn Tyr Asn Arg Arg Phe Thr Pro
115 120 125
Glu Ile Ala Glu Arg Pro Lys Val Arg Asp Gln Ala Gly Arg Met Asn
130 135 140
Tyr Tyr Trp Thr Leu Leu Lys Pro Gly Asp Thr Ile Ile Phe Glu Ala
145 150 155 160
Asn Gly Asn Leu Ile Ala Pro Met Tyr Ala Ala Ala Leu Ser Arg Gly
165 170 175

<210> 26
<211> 223
<212> PRT
<213> Artificial Sequence

<220>
<223> VN HA1-2mut

<400> 26
Gly Ala Lys Pro Leu Ser Leu Arg Asp Cys Ser Val Ala Gly Trp Leu
1 5 10 15
Leu Gly Asn Pro Met Cys Asp Glu Phe Ile Asn Val Pro Glu Trp Ser
20 25 30
Asp Ile Ala Glu Lys Ala Asn Pro Val Asn Asp Leu Cys Tyr Pro Gly
35 40 45
Asp Phe Asn Asp Tyr Glu Glu Leu Lys His Leu Leu Ser Arg Ile Asn
50 55 60
His Phe Glu Lys Ile Gln Ile Ile Pro Lys Ser Ser Trp Ser Ser His
65 70 75 80
Glu Ala Ser Leu Gly Val Ser Ser Ala Cys Pro Tyr Gln Gly Lys Ser
85 90 95
Ser Phe Phe Arg Asn Val Val Trp Leu Ile Lys Lys Asn Ser Thr Tyr
100 105 110

```
Pro Thr Ile Lys Arg Ser Tyr Asn Asn Thr Asn Gln Glu Asp Leu Leu
        115                 120                 125
Val Leu Trp Gly Ile His His Pro Asn Asp Ala Ala Glu Gln Thr Lys
    130                 135                 140
Leu Tyr Gln Asn Pro Thr Thr Tyr Ile Ser Val Gly Thr Ser Thr Leu
145                 150                 155                 160
Asn Gln Arg Leu Val Pro Arg Ile Ala Thr Arg Ser Lys Val Asn Gly
                165                 170                 175
Gln Ser Gly Arg Met Glu Phe Phe Trp Thr Ile Leu Lys Pro Asn Asp
            180                 185                 190
Ala Ile Asn Phe Glu Ser Asn Gly Asn Phe Ile Ala Pro Glu Tyr Ala
        195                 200                 205
Tyr Lys Ile Val Lys Lys Gly Asp Ser Thr Ile Met Lys Ser Glu
    210                 215                 220


<210> 27
<211> 174
<212> PRT
<213> Artificial Sequence

<220>
<223> VN HA1-3mut

<400> 27
Asn Asp Leu Cys Tyr Pro Gly Asp Phe Asn Asp Lys Glu Glu Leu Lys
 1               5                  10                  15
His Leu Leu Ser Arg Ile Asn His Phe Glu Lys Ile Gln Ile Ile Pro
            20                  25                  30
Lys Ser Ser Trp Ser Ser His Glu Ala Ser Leu Gly Val Ser Ser Ala
        35                  40                  45
Cys Pro Tyr Gln Gly Lys Ser Ser Phe Phe Arg Asn Val Val Trp Leu
    50                  55                  60
Ile Lys Lys Asn Ser Thr Tyr Pro Thr Ile Lys Arg Ser Tyr Asn Asn
65                  70                  75                  80
Thr Asn Gln Glu Asp Leu Leu Val Leu Trp Gly Ile His His Pro Asn
                85                  90                  95
Asp Ala Ala Glu Gln Thr Lys Leu Tyr Gln Asn Pro Thr Thr Tyr Ile
            100                 105                 110
Ser Val Gly Thr Ser Thr Leu Asn Gln Arg Leu Val Pro Arg Ile Ala
        115                 120                 125
Thr Arg Ser Lys Val Asn Gly Gln Ser Gly Arg Met Glu Phe Phe Trp
    130                 135                 140
Thr Ile Leu Lys Pro Asn Asp Ala Ile Asn Phe Glu Ser Asn Gly Asn
145                 150                 155                 160
Phe Ile Ala Pro Glu Tyr Ala Tyr Lys Ile Ala Lys Lys Gly
                165                 170


<210> 28
<211> 223
<212> PRT
<213> Artificial Sequence

<220>
<223> IND HA1-2mut

<400> 28
Gly Ala Lys Pro Leu Ser Leu Arg Asp Cys Ser Val Ala Gly Trp Leu
```

```
 1               5                  10                  15
Leu Gly Asn Pro Met Cys Asp Glu Phe Ile Asn Val Pro Glu Trp Ser
            20                  25                  30
Asp Ile Ala Glu Lys Ala Asn Pro Thr Asn Asp Leu Cys Tyr Pro Gly
        35                  40                  45
Ser Phe Asn Asp Tyr Glu Glu Leu Lys His Leu Leu Ser Arg Ile Asn
    50                  55                  60
His Phe Glu Lys Ile Gln Ile Ile Pro Lys Ser Ser Trp Ser Asp His
65                  70                  75                  80
Glu Ala Ser Ser Gly Val Ser Ser Ala Cys Pro Tyr Leu Gly Ser Pro
                85                  90                  95
Ser Phe Phe Arg Asn Val Val Trp Leu Ile Lys Lys Asn Ser Thr Tyr
            100                 105                 110
Pro Thr Ile Lys Lys Ser Tyr Asn Asn Thr Asn Gln Glu Asp Leu Leu
        115                 120                 125
Val Leu Trp Gly Ile His His Pro Asn Asp Ala Ala Glu Gln Thr Arg
    130                 135                 140
Leu Tyr Gln Asn Pro Thr Thr Tyr Ile Ser Ile Gly Thr Ser Thr Leu
145                 150                 155                 160
Asn Gln Arg Leu Val Pro Lys Ile Ala Thr Arg Ser Lys Val Asn Gly
                165                 170                 175
Gln Ser Gly Arg Met Glu Phe Phe Trp Thr Ile Leu Lys Pro Asn Asp
            180                 185                 190
Ala Ile Asn Phe Glu Ser Asn Gly Asn Phe Ile Ala Pro Glu Tyr Ala
        195                 200                 205
Tyr Lys Ile Val Lys Lys Gly Asp Ser Ala Ile Met Lys Ser Glu
    210                 215                 220
```

<210> 29
<211> 174
<212> PRT
<213> Artificial Sequence

<220>
<223> IND HA1-3mut

<400> 29

```
Asn Asp Leu Cys Tyr Pro Gly Ser Phe Asn Asp Lys Glu Glu Leu Lys
 1               5                  10                  15
His Leu Leu Ser Arg Ile Asn His Phe Glu Lys Ile Gln Ile Ile Pro
            20                  25                  30
Lys Ser Ser Trp Ser Asp His Glu Ala Ser Ser Gly Val Ser Ser Ala
        35                  40                  45
Cys Pro Tyr Leu Gly Ser Pro Ser Phe Phe Arg Asn Val Val Trp Leu
    50                  55                  60
Ile Lys Lys Asn Ser Thr Tyr Pro Thr Ile Lys Lys Ser Tyr Asn Asn
65                  70                  75                  80
Thr Asn Gln Glu Asp Leu Leu Val Leu Trp Gly Ile His His Pro Asn
                85                  90                  95
Asp Ala Ala Glu Gln Thr Arg Leu Tyr Gln Asn Pro Thr Thr Tyr Ile
            100                 105                 110
Ser Ile Gly Thr Ser Thr Leu Asn Gln Arg Leu Val Pro Lys Ile Ala
        115                 120                 125
Thr Arg Ser Lys Val Asn Gly Gln Ser Gly Arg Met Glu Phe Phe Trp
    130                 135                 140
Thr Ile Leu Lys Pro Asn Asp Ala Ile Asn Phe Glu Ser Asn Gly Asn
145                 150                 155                 160
Phe Ile Ala Pro Glu Tyr Ala Tyr Lys Ile Ala Lys Lys Gly
```

165 170

<210> 30
<211> 223
<212> PRT
<213> Artificial Sequence

<220>
<223> NC HA1-2mut

<400> 30
Lys Gly Ala Ala Pro Leu Gln Leu Gly Asn Cys Ser Val Ala Gly Trp
1 5 10 15
Ile Leu Gly Asn Pro Glu Cys Glu Leu Leu Ile Ser Lys Glu Ser Trp
20 25 30
Ser Asp Ile Ala Glu Thr Pro Asn Pro Glu Asn Gly Thr Cys Tyr Pro
35 40 45
Gly Tyr Phe Ala Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val
50 55 60
Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Glu Ser Ser Trp Pro
65 70 75 80
Asn His Thr Val Thr Gly Val Ser Ala Ser Cys Ser His Asn Gly Lys
85 90 95
Ser Ser Phe Tyr Arg Asn Leu Leu Trp Leu Thr Gly Lys Asn Gly Leu
100 105 110
Tyr Pro Asn Leu Ser Lys Ser Tyr Val Asn Asn Lys Glu Lys Glu Val
115 120 125
Leu Val Leu Trp Gly Val His His Pro Pro Asn Ile Gly Asn Gln Arg
130 135 140
Ala Leu Tyr His Thr Glu Asn Ala Tyr Val Ser Val Val Ser Ser His
145 150 155 160
Tyr Ser Arg Arg Phe Thr Pro Glu Ile Ala Lys Arg Pro Lys Val Arg
165 170 175
Asp Gln Glu Gly Arg Ile Asn Tyr Tyr Trp Thr Leu Leu Glu Pro Gly
180 185 190
Asp Thr Ile Ile Phe Glu Ala Asn Gly Asn Leu Ile Ala Pro Trp Tyr
195 200 205
Ala Phe Ala Leu Ser Arg Gly Phe Gly Ser Gly Ile Ile Thr Ser
210 215 220

<210> 31
<211> 176
<212> PRT
<213> Artificial Sequence

<220>
<223> NC HA1-3mut

<400> 31
Asn Pro Glu Asn Glu Thr Cys Tyr Pro Gly Tyr Phe Ala Asp Lys Glu
1 5 10 15
Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe Glu Arg Phe Glu
20 25 30
Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His Thr Val Thr Gly Val
35 40 45
Ser Ala Ser Cys Ser His Asn Gly Lys Ser Ser Phe Tyr Arg Asn Leu
50 55 60

```
Leu Trp Leu Thr Gly Lys Asn Gly Leu Tyr Pro Asn Leu Ser Lys Ser
65                  70                  75                  80
Tyr Val Asn Asn Lys Glu Lys Glu Val Leu Val Leu Trp Gly Val His
                85                  90                  95
His Pro Pro Asn Ile Gly Asn Gln Arg Ala Leu Tyr His Thr Glu Asn
            100                 105                 110
Ala Tyr Val Ser Val Val Ser Ser His Tyr Ser Arg Arg Phe Thr Pro
        115                 120                 125
Glu Ile Ala Lys Arg Pro Lys Val Arg Asp Gln Glu Gly Arg Ile Asn
    130                 135                 140
Tyr Tyr Trp Thr Leu Leu Glu Pro Gly Asp Thr Ile Ile Phe Glu Ala
145                 150                 155                 160
Asn Gly Asn Leu Ile Ala Pro Trp Tyr Ala Ala Ala Leu Ser Arg Gly
                165                 170                 175


<210> 32
<211> 218
<212> PRT
<213> Artificial Sequence

<220>
<223> WIS HA1-2mut

<400> 32
Ser Pro His Gln Ala Leu Asp Gly Glu Asn Cys Thr Leu Ile Asp Ala
 1               5                  10                  15
Leu Leu Gly Asp Pro Gln Cys Asp Gly Phe Gln Asn Lys Lys Trp Asp
            20                  25                  30
Asp Phe Ala Glu Arg Ser Lys Ala Tyr Ser Asn Cys Tyr Pro Tyr Asp
        35                  40                  45
Val Pro Asp Tyr Ala Ser Leu Arg Ser Leu Val Ala Ser Ser Gly Thr
    50                  55                  60
Leu Glu Phe Asn Asp Glu Ser Phe Asn Trp Thr Gly Val Thr Gln Asn
65                  70                  75                  80
Gly Thr Ser Ser Ala Cys Lys Arg Arg Ser Asn Asn Ser Phe Phe Ser
                85                  90                  95
Arg Leu Asn Trp Leu Thr His Leu Lys Phe Lys Tyr Pro Ala Leu Asn
            100                 105                 110
Val Thr Met Pro Asn Asn Glu Lys Phe Asp Lys Leu Tyr Ile Trp Gly
        115                 120                 125
Val His His Pro Gly Thr Asp Asn Asp Gln Ile Phe Leu His Ala Gln
    130                 135                 140
Ala Ser Gly Arg Ile Thr Val Ser Thr Lys Arg Ser Gln Gln Thr Val
145                 150                 155                 160
Ile Pro Asn Ile Gly Ser Arg Pro Arg Ile Arg Asn Ile Pro Ser Arg
                165                 170                 175
Ile Ser Ile Tyr Trp Thr Ile Val Lys Pro Gly Asp Ile Leu Leu Ile
            180                 185                 190
Asn Ser Thr Gly Asn Leu Ile Ala Pro Arg Gly Tyr Phe Lys Ile Arg
        195                 200                 205
Ser Gly Lys Ser Ser Ile Met Arg Ser Asp
    210                 215


<210> 33
<211> 169
<212> PRT
<213> Artificial Sequence
```

<220>
<223> WIS HA1-3mut

<400> 33
Ser Asn Cys Tyr Pro Lys Asp Ser Pro Asp Glu Ala Ser Leu Arg Ser
1 5 10 15
Leu Val Ala Ser Ser Gly Thr Asp Glu Phe Asn Asp Glu Ser Phe Asn
20 25 30
Trp Thr Gly Val Thr Gln Asn Gly Thr Ser Ser Ala Cys Lys Arg Arg
35 40 45
Ser Asn Asn Ser Phe Phe Ser Arg Leu Asn Trp Leu Thr His Leu Lys
50 55 60
Phe Lys Tyr Pro Ala Leu Asn Val Thr Met Pro Asn Asn Glu Lys Phe
65 70 75 80
Asp Lys Leu Tyr Ile Trp Gly Val His His Pro Gly Thr Asp Asn Asp
85 90 95
Gln Ile Phe Leu His Ala Gln Ala Ser Gly Arg Ile Thr Val Ser Thr
100 105 110
Lys Arg Ser Gln Gln Thr Val Ile Pro Asn Ile Gly Ser Arg Pro Arg
115 120 125
Ile Arg Asn Ile Pro Ser Arg Ile Ser Ile Tyr Trp Thr Ile Val Lys
130 135 140
Pro Gly Asp Ile Leu Leu Ile Asn Ser Thr Gly Asn Leu Ile Ala Pro
145 150 155 160
Arg Gly Tyr Phe Lys Ile Arg Ser Gly
165

<210> 34
<211> 320
<212> PRT
<213> Influenza A/Swine/Hong Kong/9/98 HA

<400> 34
Asp Lys Ile Cys Ile Gly Tyr Gln Ser Thr Asn Ser Thr Glu Thr Val
1 5 10 15
Asp Thr Leu Thr Glu Thr Asn Val Pro Val Thr His Ala Lys Glu Leu
20 25 30
Leu His Thr Glu His Asn Gly Met Leu Cys Ala Thr Asn Leu Gly His
35 40 45
Pro Leu Ile Leu Asp Thr Cys Thr Ile Glu Gly Leu Ile Tyr Gly Asn
50 55 60
Pro Ser Cys Asp Leu Leu Leu Gly Gly Arg Glu Trp Ser Tyr Ile Val
65 70 75 80
Glu Arg Pro Ser Ala Val Asn Gly Met Cys Tyr Pro Gly Asn Val Glu
85 90 95
Asn Leu Glu Glu Leu Arg Ser Leu Phe Ser Ser Ala Ser Ser Tyr Gln
100 105 110
Arg Ile Gln Ile Phe Pro Asp Thr Ile Trp Asn Val Ser Tyr Ser Gly
115 120 125
Thr Ser Lys Ala Cys Ser Asp Ser Phe Tyr Arg Ser Met Arg Trp Leu
130 135 140
Thr Gln Lys Asn Asn Ala Tyr Pro Ile Gln Asp Ala Gln Tyr Thr Asn
145 150 155 160
Asn Arg Gly Lys Ser Ile Leu Phe Met Trp Gly Ile Asn His Pro Pro
165 170 175
Thr Asp Thr Val Gln Thr Asn Leu Tyr Thr Arg Thr Asp Thr Thr Thr
180 185 190

```
Ser Val Thr Thr Glu Asp Ile Asn Arg Thr Phe Lys Pro Val Ile Gly
        195                 200                 205
Pro Arg Pro Leu Val Asn Gly Leu His Gly Arg Ile Asp Tyr Tyr Trp
    210                 215                 220
Ser Val Leu Lys Pro Gly Gln Thr Leu Arg Val Arg Ser Asn Gly Asn
225                 230                 235                 240
Leu Ile Ala Pro Trp Tyr Gly His Ile Leu Ser Gly Glu Ser His Gly
                245                 250                 255
Arg Ile Leu Lys Thr Asp Leu Asn Ser Gly Asn Cys Val Val Gln Cys
            260                 265                 270
Gln Thr Glu Arg Gly Gly Leu Asn Thr Thr Leu Pro Phe His Asn Val
        275                 280                 285
Ser Lys Tyr Ala Phe Gly Asn Cys Pro Lys Tyr Val Gly Val Lys Ser
    290                 295                 300
Leu Lys Leu Ala Val Gly Leu Arg Asn Val Pro Ala Arg Ser Ser Arg
305                 310                 315                 320


<210> 35
<211> 566
<212> PRT
<213> Influenza A//Aichi/2/1968 HA

<400> 35
Met Lys Thr Ile Ile Ala Leu Ser Tyr Ile Phe Cys Leu Ala Leu Gly
 1                   5                  10                  15
Gln Asp Leu Pro Gly Asn Asp Asn Ser Thr Ala Thr Leu Cys Leu Gly
            20                  25                  30
His His Ala Val Pro Asn Gly Thr Leu Val Lys Thr Ile Thr Asp Asp
        35                  40                  45
Gln Ile Glu Val Thr Asn Ala Thr Glu Leu Val Gln Ser Ser Ser Thr
    50                  55                  60
Gly Lys Ile Cys Asn Asn Pro His Arg Ile Leu Asp Gly Ile Asp Cys
65                  70                  75                  80
Thr Leu Ile Asp Ala Leu Leu Gly Asp Pro His Cys Asp Val Phe Gln
                85                  90                  95
Asn Glu Thr Trp Asp Leu Phe Val Glu Arg Ser Lys Ala Phe Ser Asn
            100                 105                 110
Cys Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Leu Arg Ser Leu Val
        115                 120                 125
Ala Ser Ser Gly Thr Leu Glu Phe Ile Thr Glu Gly Phe Thr Trp Thr
    130                 135                 140
Gly Val Thr Gln Asn Gly Gly Ser Asn Ala Cys Lys Arg Gly Pro Gly
145                 150                 155                 160
Ser Gly Phe Phe Ser Arg Leu Asn Trp Leu Thr Lys Ser Gly Ser Thr
                165                 170                 175
Tyr Pro Val Leu Asn Val Thr Met Pro Asn Asn Asp Asn Phe Asp Lys
            180                 185                 190
Leu Tyr Ile Trp Gly Ile His His Pro Ser Thr Asn Gln Glu Gln Thr
        195                 200                 205
Ser Leu Tyr Val Gln Ala Ser Gly Arg Val Thr Val Ser Thr Arg Arg
    210                 215                 220
Ser Gln Gln Thr Ile Ile Pro Asn Ile Gly Ser Arg Pro Trp Val Arg
225                 230                 235                 240
Gly Leu Ser Ser Arg Ile Ser Ile Tyr Trp Thr Ile Val Lys Pro Gly
                245                 250                 255
Asp Val Leu Val Ile Asn Ser Asn Gly Asn Leu Ile Ala Pro Arg Gly
            260                 265                 270
Tyr Phe Lys Met Arg Thr Gly Lys Ser Ser Ile Met Arg Ser Asp Ala
```

```
        275                 280                 285
Pro Ile Asp Thr Cys Ile Ser Glu Cys Ile Thr Pro Asn Gly Ser Ile
    290                 295                 300
Pro Asn Asp Lys Pro Phe Gln Asn Val Asn Lys Ile Thr Tyr Gly Ala
305                 310                 315                 320
Cys Pro Lys Tyr Val Lys Gln Asn Thr Leu Lys Leu Ala Thr Gly Met
                325                 330                 335
Arg Asn Val Pro Glu Lys Gln Thr Arg Gly Leu Phe Gly Ala Ile Ala
            340                 345                 350
Gly Phe Ile Glu Asn Gly Trp Glu Gly Met Ile Asp Gly Trp Tyr Gly
        355                 360                 365
Phe Arg His Gln Asn Ser Glu Gly Thr Gly Gln Ala Ala Asp Leu Lys
    370                 375                 380
Ser Thr Gln Ala Ala Ile Asp Gln Ile Asn Gly Lys Leu Asn Arg Val
385                 390                 395                 400
Ile Glu Lys Thr Asn Glu Lys Phe His Gln Ile Glu Lys Glu Phe Ser
                405                 410                 415
Glu Val Glu Gly Arg Ile Gln Asp Leu Glu Lys Tyr Val Glu Asp Thr
            420                 425                 430
Lys Ile Asp Leu Trp Ser Tyr Asn Ala Glu Leu Leu Val Ala Leu Glu
        435                 440                 445
Asn Gln His Thr Ile Asp Leu Thr Asp Ser Glu Met Asn Lys Leu Phe
    450                 455                 460
Glu Lys Thr Arg Arg Gln Leu Arg Glu Asn Ala Glu Glu Met Gly Asn
465                 470                 475                 480
Gly Cys Phe Lys Ile Tyr His Lys Cys Asp Asn Ala Cys Ile Glu Ser
                485                 490                 495
Ile Arg Asn Gly Thr Tyr Asp His Asp Val Tyr Arg Asp Glu Ala Leu
            500                 505                 510
Asn Asn Arg Phe Gln Ile Lys Gly Val Glu Leu Lys Ser Gly Tyr Lys
        515                 520                 525
Asp Trp Ile Leu Trp Ile Ser Phe Ala Ile Ser Cys Phe Leu Leu Cys
    530                 535                 540
Val Val Leu Leu Gly Phe Ile Met Trp Ala Cys Gln Arg Gly Asn Ile
545                 550                 555                 560
Arg Cys Asn Ile Cys Ile
                565


<210> 36
<211> 584
<212> PRT
<213> Influenza B/Lee/40 HA

<400> 36
Met Lys Ala Ile Ile Val Leu Leu Met Val Val Thr Ser Asn Ala Asp
 1               5                  10                  15
Arg Ile Cys Thr Gly Ile Thr Ser Ser Asn Ser Pro His Val Val Lys
            20                  25                  30
Thr Ala Thr Gln Gly Glu Val Asn Val Thr Gly Val Ile Pro Leu Thr
        35                  40                  45
Thr Thr Pro Thr Lys Ser His Phe Ala Asn Leu Lys Gly Thr Gln Thr
    50                  55                  60
Arg Gly Lys Leu Cys Pro Asn Cys Phe Asn Cys Thr Asp Leu Asp Val
65                  70                  75                  80
Ala Leu Gly Arg Pro Lys Cys Met Gly Asn Thr Pro Ser Ala Lys Val
                85                  90                  95
Ser Ile Leu His Glu Val Lys Pro Ala Thr Ser Gly Cys Phe Pro Ile
            100                 105                 110
```

```
Met His Asp Arg Thr Lys Ile Arg Gln Leu Pro Asn Leu Leu Arg Gly
        115                 120                 125
Tyr Glu Asn Ile Arg Leu Ser Thr Ser Asn Val Ile Asn Thr Glu Thr
    130                 135                 140
Ala Pro Gly Gly Pro Tyr Lys Val Gly Thr Ser Gly Ser Cys Pro Asn
145                 150                 155                 160
Val Ala Asn Gly Asn Gly Phe Phe Asn Thr Met Ala Trp Val Ile Pro
                165                 170                 175
Lys Asp Asn Asn Lys Thr Ala Ile Asn Pro Val Thr Val Glu Val Pro
            180                 185                 190
Tyr Ile Cys Ser Glu Gly Glu Asp Gln Ile Thr Val Trp Gly Phe His
        195                 200                 205
Ser Asp Asp Lys Thr Gln Met Glu Arg Leu Tyr Gly Asp Ser Asn Pro
    210                 215                 220
Gln Lys Phe Thr Ser Ser Ala Asn Gly Val Thr Thr His Tyr Val Ser
225                 230                 235                 240
Gln Ile Gly Gly Phe Pro Asn Gln Thr Glu Asp Glu Gly Leu Lys Gln
                245                 250                 255
Ser Gly Arg Ile Val Val Asp Tyr Met Val Gln Lys Pro Gly Lys Thr
            260                 265                 270
Gly Thr Ile Val Tyr Gln Arg Gly Ile Leu Leu Pro Gln Lys Val Trp
        275                 280                 285
Cys Ala Ser Gly Arg Ser Lys Val Ile Lys Gly Ser Leu Pro Leu Ile
    290                 295                 300
Gly Glu Ala Asp Cys Leu His Glu Lys Tyr Gly Gly Leu Asn Lys Ser
305                 310                 315                 320
Lys Pro Tyr Tyr Thr Gly Glu His Ala Lys Ala Ile Gly Asn Cys Pro
                325                 330                 335
Ile Trp Val Lys Thr Pro Leu Lys Leu Ala Asn Gly Thr Lys Tyr Arg
            340                 345                 350
Pro Pro Ala Lys Leu Leu Lys Glu Arg Gly Phe Phe Gly Ala Ile Ala
        355                 360                 365
Gly Phe Leu Glu Gly Gly Trp Glu Gly Met Ile Ala Gly Trp His Gly
    370                 375                 380
Tyr Thr Ser His Gly Ala His Gly Val Ala Val Ala Ala Asp Leu Lys
385                 390                 395                 400
Ser Thr Gln Glu Ala Ile Asn Lys Ile Thr Lys Asn Leu Asn Tyr Leu
                405                 410                 415
Ser Glu Leu Glu Val Lys Asn Leu Gln Arg Leu Ser Gly Ala Met Asn
            420                 425                 430
Glu Leu His Asp Glu Ile Leu Glu Leu Asp Glu Lys Val Asp Asp Leu
        435                 440                 445
Arg Ala Asp Thr Ile Ser Ser Gln Ile Glu Leu Ala Val Leu Leu Ser
    450                 455                 460
Asn Glu Gly Ile Ile Asn Ser Glu Asp Glu His Leu Leu Ala Leu Glu
465                 470                 475                 480
Arg Lys Leu Lys Lys Met Leu Gly Pro Ser Ala Val Glu Ile Gly Asn
                485                 490                 495
Gly Cys Phe Glu Thr Lys His Lys Cys Asn Gln Thr Cys Leu Asp Arg
            500                 505                 510
Ile Ala Ala Gly Thr Phe Asn Ala Gly Asp Phe Ser Leu Pro Thr Phe
        515                 520                 525
Asp Ser Leu Asn Ile Thr Ala Ala Ser Leu Asn Asp Asp Gly Leu Asp
    530                 535                 540
Asn His Thr Ile Leu Leu Tyr Tyr Ser Thr Ala Ala Ser Ser Leu Ala
545                 550                 555                 560
Val Thr Leu Met Ile Ala Ile Phe Ile Val Tyr Met Val Ser Arg Asp
                565                 570                 575
Asn Val Ser Cys Ser Ile Cys Leu
```

580

<210> 37
<211> 655
<212> PRT
<213> Influenza C/Johannesburg/1/66

<400> 37

Met	Phe	Phe	Ser	Leu	Leu	Leu	Val	Leu	Gly	Leu	Thr	Glu	Ala	Glu	Lys
1				5					10					15	
Ile	Lys	Ile	Cys	Leu	Gln	Lys	Gln	Val	Asn	Ser	Ser	Phe	Ser	Leu	His
			20					25					30		
Asn	Gly	Phe	Gly	Gly	Asn	Leu	Tyr	Ala	Thr	Glu	Glu	Lys	Arg	Met	Phe
		35					40					45			
Glu	Leu	Val	Lys	Pro	Lys	Ala	Gly	Ala	Ser	Val	Leu	Asn	Gln	Ser	Thr
	50					55					60				
Trp	Ile	Gly	Phe	Gly	Asp	Ser	Arg	Thr	Asp	Lys	Ser	Asn	Ser	Ala	Phe
65					70					75					80
Pro	Arg	Ser	Ala	Asp	Val	Ser	Ala	Lys	Thr	Ala	Asp	Lys	Phe	Arg	Ser
				85					90					95	
Leu	Ser	Gly	Gly	Ser	Leu	Met	Leu	Ser	Met	Phe	Gly	Pro	Pro	Gly	Lys
			100					105					110		
Val	Asp	Tyr	Leu	Tyr	Gln	Gly	Cys	Gly	Lys	His	Lys	Val	Phe	Tyr	Glu
		115					120					125			
Gly	Val	Asn	Trp	Ser	Pro	His	Ala	Ala	Ile	Asn	Cys	Tyr	Arg	Lys	Asn
	130					135					140				
Trp	Thr	Asp	Ile	Lys	Leu	Asn	Phe	Gln	Lys	Asn	Ile	Tyr	Glu	Leu	Ala
145					150					155					160
Ser	Gln	Ser	His	Cys	Met	Ser	Leu	Val	Asn	Ala	Leu	Asp	Lys	Thr	Ile
				165					170					175	
Pro	Leu	Gln	Val	Thr	Ala	Gly	Thr	Ala	Gly	Asn	Cys	Asn	Asn	Ser	Phe
			180					185					190		
Leu	Lys	Asn	Pro	Ala	Leu	Tyr	Thr	Gln	Glu	Val	Lys	Pro	Ser	Glu	Asn
		195					200					205			
Lys	Cys	Gly	Lys	Glu	Asn	Leu	Ala	Phe	Phe	Thr	Leu	Pro	Thr	Gln	Phe
	210					215					220				
Gly	Thr	Tyr	Glu	Cys	Lys	Leu	His	Leu	Val	Ala	Ser	Cys	Tyr	Phe	Ile
225					230					235					240
Tyr	Asp	Ser	Lys	Glu	Val	Tyr	Asn	Lys	Arg	Gly	Cys	Asp	Asn	Tyr	Phe
				245					250					255	
Gln	Val	Ile	Tyr	Asp	Ser	Ser	Gly	Lys	Val	Val	Gly	Gly	Leu	Asp	Asn
			260					265					270		
Arg	Val	Ser	Pro	Tyr	Thr	Gly	Asn	Ser	Gly	Asp	Thr	Pro	Thr	Met	Gln
		275					280					285			
Cys	Asp	Met	Leu	Gln	Leu	Lys	Pro	Gly	Arg	Tyr	Ser	Val	Arg	Ser	Ser
	290					295					300				
Pro	Arg	Phe	Leu	Leu	Met	Pro	Glu	Arg	Ser	Tyr	Cys	Phe	Asp	Met	Lys
305					310					315					320
Glu	Lys	Gly	Pro	Val	Thr	Ala	Val	Gln	Ser	Ile	Trp	Gly	Lys	Gly	Arg
				325					330					335	
Glu	Ser	Asp	Tyr	Ala	Val	Asp	Gln	Ala	Cys	Leu	Ser	Thr	Pro	Gly	Cys
			340					345					350		
Met	Leu	Ile	Gln	Lys	Gln	Lys	Pro	Tyr	Ile	Gly	Glu	Ala	Asp	Asp	His
		355					360					365			
His	Gly	Asp	Gln	Glu	Met	Arg	Glu	Leu	Leu	Ser	Gly	Leu	Asp	Tyr	Glu
	370					375					380				
Ala	Arg	Cys	Ile	Ser	Gln	Ser	Gly	Trp	Val	Asn	Glu	Thr	Ser	Pro	Phe
385					390					395					400

```
Thr Glu Lys Tyr Leu Leu Pro Pro Lys Phe Gly Arg Cys Pro Leu Ala
                405                 410                 415
Ala Lys Glu Glu Ser Ile Pro Lys Ile Pro Asp Gly Leu Leu Ile Pro
            420                 425                 430
Thr Ser Gly Thr Asp Thr Thr Val Thr Lys Pro Lys Ser Arg Ile Phe
        435                 440                 445
Gly Ile Asp Asp Leu Ile Ile Gly Leu Leu Phe Val Ala Ile Val Glu
    450                 455                 460
Ala Gly Ile Gly Gly Tyr Leu Leu Gly Ser Arg Lys Glu Ser Gly Gly
465                 470                 475                 480
Gly Val Thr Lys Glu Ser Ala Glu Lys Gly Phe Glu Lys Ile Gly Asn
                485                 490                 495
Asp Ile Gln Ile Leu Lys Ser Ser Ile Asn Ile Ala Ile Glu Lys Leu
            500                 505                 510
Asn Asp Arg Ile Ser His Asp Glu Gln Ala Ile Arg Asp Leu Thr Leu
        515                 520                 525
Glu Ile Glu Asn Ala Arg Ser Glu Ala Leu Leu Gly Glu Leu Gly Ile
    530                 535                 540
Ile Arg Ala Leu Leu Val Gly Asn Ile Ser Ile Gly Leu Gln Glu Ser
545                 550                 555                 560
Leu Trp Glu Leu Ala Ser Glu Ile Thr Asn Arg Ala Gly Asp Leu Ala
                565                 570                 575
Val Glu Val Ser Pro Gly Cys Trp Ile Ile Asp Asn Asn Ile Cys Asp
            580                 585                 590
Gln Ser Cys Gln Asn Phe Ile Phe Lys Phe Asn Glu Thr Ala Pro Val
        595                 600                 605
Pro Thr Ile Pro Pro Leu Asp Thr Lys Ile Asp Leu Gln Ser Asp Pro
    610                 615                 620
Phe Tyr Trp Gly Ser Ser Leu Gly Leu Ala Ile Thr Ala Thr Ile Ser
625                 630                 635                 640
Leu Ala Ala Leu Val Ile Ser Gly Ile Ala Ile Cys Arg Thr Lys
                645                 650                 655


<210> 38
<211> 293
<212> PRT
<213> Artificial Sequence

<220>
<223> B/Lee/40 HA1-1

<400> 38
Thr Thr Thr Pro Thr Lys Ser His Phe Ala Asn Leu Lys Gly Thr Gln
 1               5                  10                  15
Thr Arg Gly Lys Leu Cys Pro Asn Cys Phe Asn Cys Thr Asp Leu Asp
            20                  25                  30
Val Ala Leu Gly Arg Pro Lys Cys Met Gly Asn Thr Pro Ser Ala Lys
        35                  40                  45
Val Ser Ile Leu His Glu Val Lys Pro Ala Thr Ser Gly Cys Phe Pro
    50                  55                  60
Ile Met His Asp Arg Thr Lys Ile Arg Gln Leu Pro Asn Leu Leu Arg
65                  70                  75                  80
Gly Tyr Glu Asn Ile Arg Leu Ser Thr Ser Asn Val Ile Asn Thr Glu
                85                  90                  95
Thr Ala Pro Gly Gly Pro Tyr Lys Val Gly Thr Ser Gly Ser Cys Pro
            100                 105                 110
Asn Val Ala Asn Gly Asn Gly Phe Phe Asn Thr Met Ala Trp Val Ile
        115                 120                 125
```

```
Pro Lys Asp Asn Asn Lys Thr Ala Ile Asn Pro Val Thr Val Glu Val
    130                 135                 140
Pro Tyr Ile Cys Ser Glu Gly Glu Asp Gln Ile Thr Val Trp Gly Phe
145                 150                 155                 160
His Ser Asp Asp Lys Thr Gln Met Glu Arg Leu Tyr Gly Asp Ser Asn
                165                 170                 175
Pro Gln Lys Phe Thr Ser Ser Ala Asn Gly Val Thr Thr His Tyr Val
            180                 185                 190
Ser Gln Ile Gly Gly Phe Pro Asn Gln Thr Glu Asp Glu Gly Leu Lys
        195                 200                 205
Gln Ser Gly Arg Ile Val Val Asp Tyr Met Val Gln Lys Pro Gly Lys
    210                 215                 220
Thr Gly Thr Ile Val Tyr Gln Arg Gly Ile Leu Leu Pro Gln Lys Val
225                 230                 235                 240
Trp Cys Ala Ser Gly Arg Ser Lys Val Ile Lys Gly Ser Leu Pro Leu
                245                 250                 255
Ile Gly Glu Ala Asp Cys Leu His Glu Lys Tyr Gly Gly Leu Asn Lys
            260                 265                 270
Ser Lys Pro Tyr Tyr Thr Gly Glu His Ala Lys Ala Ile Gly Asn Cys
        275                 280                 285
Pro Ile Trp Val Lys
    290
```

<210> 39
<211> 240
<212> PRT
<213> Artificial Sequence

<220>
<223> B/Lee/40 HA1-2

```
<400> 39
Lys Gly Thr Gln Thr Arg Gly Lys Leu Cys Pro Asn Cys Phe Asn Cys
 1               5                  10                  15
Thr Asp Leu Asp Val Ala Leu Gly Arg Pro Lys Cys Met Gly Asn Thr
            20                  25                  30
Pro Ser Ala Lys Val Ser Ile Leu His Glu Val Lys Pro Ala Thr Ser
        35                  40                  45
Gly Cys Phe Pro Ile Met His Asp Arg Thr Lys Ile Arg Gln Leu Pro
    50                  55                  60
Asn Leu Leu Arg Gly Tyr Glu Asn Ile Arg Leu Ser Thr Ser Asn Val
65                  70                  75                  80
Ile Asn Thr Glu Thr Ala Pro Gly Gly Pro Tyr Lys Val Gly Thr Ser
                85                  90                  95
Gly Ser Cys Pro Asn Val Ala Asn Gly Asn Gly Phe Phe Asn Thr Met
            100                 105                 110
Ala Trp Val Ile Pro Lys Asp Asn Asn Lys Thr Ala Ile Asn Pro Val
        115                 120                 125
Thr Val Glu Val Pro Tyr Ile Cys Ser Glu Gly Glu Asp Gln Ile Thr
    130                 135                 140
Val Trp Gly Phe His Ser Asp Asp Lys Thr Gln Met Glu Arg Leu Tyr
145                 150                 155                 160
Gly Asp Ser Asn Pro Gln Lys Phe Thr Ser Ser Ala Asn Gly Val Thr
                165                 170                 175
Thr His Tyr Val Ser Gln Ile Gly Gly Phe Pro Asn Gln Thr Glu Asp
            180                 185                 190
Glu Gly Leu Lys Gln Ser Gly Arg Ile Val Val Asp Tyr Met Val Gln
        195                 200                 205
```

```
Lys Pro Gly Lys Thr Gly Thr Ile Val Tyr Gln Arg Gly Ile Leu Leu
    210                 215                 220
Pro Gln Lys Val Trp Cys Ala Ser Gly Arg Ser Lys Val Ile Lys Gly
225                 230                 235                 240


<210> 40
<211> 267
<212> PRT
<213> Artificial Sequence

<220>
<223> A/Aichi/2/68 HA1-1

<400> 40
Gln Ser Ser Ser Thr Gly Lys Ile Cys Asn Asn Pro His Arg Ile Leu
 1               5                  10                  15
Asp Gly Ile Asp Cys Thr Leu Ile Asp Ala Leu Leu Gly Asp Pro His
            20                  25                  30
Cys Asp Val Phe Gln Asn Glu Thr Trp Asp Leu Phe Val Glu Arg Ser
        35                  40                  45
Lys Ala Phe Ser Asn Cys Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser
    50                  55                  60
Leu Arg Ser Leu Val Ala Ser Ser Gly Thr Leu Glu Phe Ile Thr Glu
65                  70                  75                  80
Gly Phe Thr Trp Thr Gly Val Thr Gln Asn Gly Gly Ser Asn Ala Cys
                85                  90                  95
Lys Arg Gly Pro Gly Ser Gly Phe Phe Ser Arg Leu Asn Trp Leu Thr
            100                 105                 110
Lys Ser Gly Ser Thr Tyr Pro Val Leu Asn Val Thr Met Pro Asn Asn
        115                 120                 125
Asp Asn Phe Asp Lys Leu Tyr Ile Trp Gly Ile His His Pro Ser Thr
    130                 135                 140
Asn Gln Glu Gln Thr Ser Leu Tyr Val Gln Ala Ser Gly Arg Val Thr
145                 150                 155                 160
Val Ser Thr Arg Arg Ser Gln Gln Thr Ile Ile Pro Asn Ile Gly Ser
                165                 170                 175
Arg Pro Trp Val Arg Gly Leu Ser Ser Arg Ile Ser Ile Tyr Trp Thr
            180                 185                 190
Ile Val Lys Pro Gly Asp Val Leu Val Ile Asn Ser Asn Gly Asn Leu
        195                 200                 205
Ile Ala Pro Arg Gly Tyr Phe Lys Met Arg Thr Gly Lys Ser Ser Ile
    210                 215                 220
Met Arg Ser Asp Ala Pro Ile Asp Thr Cys Ile Ser Glu Cys Ile Thr
225                 230                 235                 240
Pro Asn Gly Ser Ile Pro Asn Asp Lys Pro Phe Gln Asn Val Asn Lys
                245                 250                 255
Ile Thr Tyr Gly Ala Cys Pro Lys Tyr Val Lys
            260                 265


<210> 41
<211> 216
<212> PRT
<213> Artificial Sequence

<220>
<223> A/Aichi/2/68 HA1-2
```

<400> 41
His Arg Ile Leu Asp Gly Ile Asp Cys Thr Leu Ile Asp Ala Leu Leu
1 5 10 15
Gly Asp Pro His Cys Asp Val Phe Gln Asn Glu Thr Trp Asp Leu Phe
20 25 30
Val Glu Arg Ser Lys Ala Phe Ser Asn Cys Tyr Pro Tyr Asp Val Pro
35 40 45
Asp Tyr Ala Ser Leu Arg Ser Leu Val Ala Ser Ser Gly Thr Leu Glu
50 55 60
Phe Ile Thr Glu Gly Phe Thr Trp Thr Gly Val Thr Gln Asn Gly Gly
65 70 75 80
Ser Asn Ala Cys Lys Arg Gly Pro Gly Ser Gly Phe Phe Ser Arg Leu
85 90 95
Asn Trp Leu Thr Lys Ser Gly Ser Thr Tyr Pro Val Leu Asn Val Thr
100 105 110
Met Pro Asn Asn Asp Asn Phe Asp Lys Leu Tyr Ile Trp Gly Ile His
115 120 125
His Pro Ser Thr Asn Gln Glu Gln Thr Ser Leu Tyr Val Gln Ala Ser
130 135 140
Gly Arg Val Thr Val Ser Thr Arg Arg Ser Gln Gln Thr Ile Ile Pro
145 150 155 160
Asn Ile Gly Ser Arg Pro Trp Val Arg Gly Leu Ser Ser Arg Ile Ser
165 170 175
Ile Tyr Trp Thr Ile Val Lys Pro Gly Asp Val Leu Val Ile Asn Ser
180 185 190
Asn Gly Asn Leu Ile Ala Pro Arg Gly Tyr Phe Lys Met Arg Thr Gly
195 200 205
Lys Ser Ser Ile Met Arg Ser Asp
210 215

<210> 42
<211> 358
<212> PRT
<213> Influenza B/Malaysia/2506/2004 HA

<400> 42
Ile Val Leu Leu Met Val Val Thr Ser Asn Ala Asp Arg Ile Cys Thr
1 5 10 15
Gly Ile Thr Ser Ser Asn Ser Pro His Val Val Lys Thr Ala Thr Gln
20 25 30
Gly Glu Val Asn Val Thr Gly Val Ile Pro Leu Thr Thr Thr Pro Thr
35 40 45
Lys Ser His Phe Ala Asn Leu Lys Gly Thr Glu Thr Arg Gly Lys Leu
50 55 60
Cys Pro Lys Cys Leu Asn Cys Thr Asp Leu Asp Val Ala Leu Gly Arg
65 70 75 80
Pro Lys Cys Thr Gly Asn Ile Pro Ser Ala Arg Val Ser Ile Leu His
85 90 95
Glu Val Arg Pro Val Thr Ser Gly Cys Phe Pro Ile Met His Asp Arg
100 105 110
Thr Lys Ile Arg Gln Leu Pro Asn Leu Leu Arg Gly Tyr Glu His Ile
115 120 125
Arg Leu Ser Thr His Asn Val Ile Asn Ala Glu Asn Ala Pro Gly Gly
130 135 140
Ser Tyr Lys Ile Gly Thr Ser Gly Ser Cys Pro Asn Val Thr Asn Gly
145 150 155 160
Asn Gly Phe Phe Ala Thr Met Ala Trp Ala Val Pro Lys Asn Asp Asn
165 170 175

Asn Lys Thr Ala Thr Asn Ser Leu Thr Ile Glu Val Pro Tyr Ile Cys
180 185 190
Thr Glu Gly Glu Asp Gln Ile Thr Val Trp Gly Phe His Ser Asp Asn
195 200 205
Glu Ala Gln Met Ala Lys Leu Tyr Gly Asp Ser Lys Pro Gln Lys Phe
210 215 220
Thr Ser Ser Ala Asn Gly Val Thr Thr His Tyr Val Ser Gln Ile Gly
225 230 235 240
Gly Phe Pro Asn Gln Thr Glu Asp Gly Gly Leu Pro Gln Ser Gly Arg
245 250 255
Ile Val Val Asp Tyr Met Val Gln Lys Ser Gly Lys Thr Gly Thr Ile
260 265 270
Thr Tyr Gln Arg Gly Ile Leu Leu Pro Gln Lys Val Trp Cys Ala Ser
275 280 285
Gly Arg Ser Lys Val Ile Lys Gly Ser Leu Pro Leu Ile Gly Glu Ala
290 295 300
Asp Cys Leu His Glu Lys Tyr Gly Gly Leu Asn Lys Ser Lys Pro Tyr
305 310 315 320
Tyr Thr Gly Glu His Ala Lys Ala Ile Gly Asn Cys Pro Ile Trp Val
325 330 335
Lys Thr Pro Leu Lys Leu Ala Asn Gly Thr Lys Tyr Arg Pro Pro Ala
340 345 350
Lys Leu Leu Lys Glu Arg
355

<210> 43
<211> 349
<212> PRT
<213> Influenza B/Ohio/1/2005 HA

<400> 43
Asp Arg Ile Cys Thr Gly Ile Thr Ser Ser Asn Ser Pro His Val Val
1 5 10 15
Lys Thr Ala Thr Gln Gly Glu Val Asn Val Thr Gly Val Ile Pro Leu
20 25 30
Thr Thr Thr Pro Thr Lys Ser His Phe Ala Asn Leu Lys Gly Thr Lys
35 40 45
Thr Arg Gly Lys Leu Cys Pro Lys Cys Leu Asn Cys Thr Asp Leu Asp
50 55 60
Val Ala Leu Gly Arg Pro Lys Cys Thr Gly Asn Ile Pro Ser Ala Glu
65 70 75 80
Val Ser Ile Leu His Glu Val Arg Pro Val Thr Ser Gly Cys Phe Pro
85 90 95
Ile Met His Asp Arg Thr Lys Ile Arg Gln Leu Pro Asn Leu Leu Arg
100 105 110
Gly Tyr Glu His Ile Arg Leu Ser Thr His Asn Val Ile Asn Ala Glu
115 120 125
Lys Ala Pro Gly Gly Pro Tyr Lys Ile Gly Thr Ser Gly Ser Cys Pro
130 135 140
Asn Val Thr Asn Gly Asn Gly Phe Phe Ala Thr Met Ala Trp Ala Val
145 150 155 160
Pro Lys Asn Asp Asn Asn Lys Thr Ala Thr Asn Ser Leu Thr Ile Glu
165 170 175
Val Pro Tyr Ile Cys Thr Glu Gly Glu Asp Gln Ile Thr Ile Trp Gly
180 185 190
Phe His Ser Asp Ser Glu Thr Gln Met Ala Lys Leu Tyr Gly Asp Ser
195 200 205
Lys Pro Gln Lys Phe Thr Ser Ser Ala Asn Gly Val Thr Thr His Tyr

```
    210                 215                 220
Val Ser Gln Ile Gly Gly Phe Pro Asn Gln Thr Glu Asp Gly Gly Leu
225                 230                 235                 240
Pro Gln Ser Gly Arg Ile Val Val Asp Tyr Met Val Gln Lys Ser Gly
                245                 250                 255
Lys Thr Gly Thr Ile Thr Tyr Gln Arg Gly Ile Leu Leu Pro Gln Lys
            260                 265                 270
Val Trp Cys Ala Ser Gly Arg Ser Lys Val Ile Lys Gly Ser Leu Pro
        275                 280                 285
Leu Ile Gly Glu Ala Asp Cys Leu His Glu Lys Tyr Gly Gly Leu Asn
    290                 295                 300
Lys Ser Lys Pro Tyr Tyr Thr Gly Glu His Ala Lys Ala Ile Gly Asn
305                 310                 315                 320
Cys Pro Ile Trp Val Lys Thr Pro Leu Lys Leu Ala Asn Gly Thr Lys
                325                 330                 335
Tyr Arg Pro Pro Ala Lys Leu Leu Lys Glu Arg Gly Phe
            340                 345


<210> 44
<211> 338
<212> PRT
<213> Influenza B/Shanghai/361/2002 HA

<400> 44
Asp Arg Ile Cys Thr Gly Ile Thr Ser Ser Asn Ser Pro His Val Val
 1               5                  10                  15
Lys Thr Ala Thr Gln Gly Glu Val Asn Val Thr Gly Val Ile Pro Leu
            20                  25                  30
Thr Thr Thr Pro Ile Lys Ser His Phe Ala Asn Leu Lys Gly Thr Arg
        35                  40                  45
Thr Arg Gly Lys Leu Cys Pro Asp Cys Leu Asn Cys Thr Asp Leu Asp
    50                  55                  60
Val Ala Leu Gly Arg Pro Met Cys Val Gly Thr Thr Pro Ser Ala Lys
65                  70                  75                  80
Ala Ser Ile Leu His Glu Val Arg Pro Val Thr Ser Gly Cys Phe Pro
                85                  90                  95
Ile Met His Asp Arg Thr Lys Ile Arg Gln Leu Pro Asn Leu Leu Arg
            100                 105                 110
Gly Tyr Glu Asn Ile Arg Leu Ser Thr Gln Asn Val Ile Asp Ala Glu
        115                 120                 125
Lys Ala Leu Gly Gly Pro Tyr Arg Leu Gly Thr Ser Gly Ser Cys Pro
    130                 135                 140
Asn Ala Thr Ser Lys Ser Gly Phe Phe Ala Thr Met Ala Trp Ala Val
145                 150                 155                 160
Pro Lys Asp Asn Asn Lys Asn Ala Thr Asn Pro Leu Thr Val Glu Val
                165                 170                 175
Pro Tyr Ile Cys Thr Glu Gly Glu Asp Gln Ile Thr Val Trp Gly Phe
            180                 185                 190
His Ser Asp Asp Lys Thr Gln Met Lys Asn Leu Tyr Gly Asp Ser Asn
        195                 200                 205
Pro Gln Lys Phe Thr Ser Ser Ala Asn Gly Val Thr Thr His Tyr Val
    210                 215                 220
Ser Gln Ile Gly Gly Phe Pro Asp Gln Thr Glu Asp Gly Gly Leu Pro
225                 230                 235                 240
Gln Ser Gly Arg Ile Val Val Asp Tyr Met Val Gln Lys Pro Gly Lys
                245                 250                 255
Thr Gly Thr Ile Val Tyr Gln Arg Gly Val Leu Leu Pro Gln Lys Val
            260                 265                 270
```

```
Trp Cys Ala Ser Gly Arg Ser Lys Val Ile Lys Gly Ser Leu Pro Leu
        275                 280                 285
Ile Gly Glu Ala Asp Cys Leu His Glu Lys Tyr Gly Gly Leu Asn Lys
    290                 295                 300
Ser Lys Pro Tyr Tyr Thr Gly Glu His Ala Lys Ala Ile Gly His Cys
305                 310                 315                 320
Pro Ile Trp Val Lys Thr Pro Leu Lys Leu Ala Asn Gly Thr Lys Tyr
                325                 330                 335
Arg Pro
```

<210> 45
<211> 294
<212> PRT
<213> Artificial Sequence

<220>
<223> B/Malaysia/2506/2004 HA1-1

<400> 45

```
Thr Thr Thr Pro Thr Lys Ser His Phe Ala Asn Leu Lys Gly Thr Glu
 1               5                  10                  15
Thr Arg Gly Lys Leu Cys Pro Lys Cys Leu Asn Cys Thr Asp Leu Asp
            20                  25                  30
Val Ala Leu Gly Arg Pro Lys Cys Thr Gly Asn Ile Pro Ser Ala Arg
        35                  40                  45
Val Ser Ile Leu His Glu Val Arg Pro Val Thr Ser Gly Cys Phe Pro
    50                  55                  60
Ile Met His Asp Arg Thr Lys Ile Arg Gln Leu Pro Asn Leu Leu Arg
65                  70                  75                  80
Gly Tyr Glu His Ile Arg Leu Ser Thr His Asn Val Ile Asn Ala Glu
                85                  90                  95
Asn Ala Pro Gly Gly Ser Tyr Lys Ile Gly Thr Ser Gly Ser Cys Pro
            100                 105                 110
Asn Val Thr Asn Gly Asn Gly Phe Phe Ala Thr Met Ala Trp Ala Val
        115                 120                 125
Pro Lys Asn Asp Asn Asn Lys Thr Ala Thr Asn Ser Leu Thr Ile Glu
    130                 135                 140
Val Pro Tyr Ile Cys Thr Glu Gly Glu Asp Gln Ile Thr Val Trp Gly
145                 150                 155                 160
Phe His Ser Asp Asn Glu Ala Gln Met Ala Lys Leu Tyr Gly Asp Ser
                165                 170                 175
Lys Pro Gln Lys Phe Thr Ser Ser Ala Asn Gly Val Thr Thr His Tyr
            180                 185                 190
Val Ser Gln Ile Gly Gly Phe Pro Asn Gln Thr Glu Asp Gly Gly Leu
        195                 200                 205
Pro Gln Ser Gly Arg Ile Val Val Asp Tyr Met Val Gln Lys Ser Gly
    210                 215                 220
Lys Thr Gly Thr Ile Thr Tyr Gln Arg Gly Ile Leu Leu Pro Gln Lys
225                 230                 235                 240
Val Trp Cys Ala Ser Gly Arg Ser Lys Val Ile Lys Gly Ser Leu Pro
                245                 250                 255
Leu Ile Gly Glu Ala Asp Cys Leu His Glu Lys Tyr Gly Gly Leu Asn
            260                 265                 270
Lys Ser Lys Pro Tyr Tyr Thr Gly Glu His Ala Lys Ala Ile Gly Asn
        275                 280                 285
Cys Pro Ile Trp Val Lys
    290
```

<210> 46
<211> 241
<212> PRT
<213> Artificial Sequence

<220>
<223> B/Malaysia/2506/2004 HA1-2

<400> 46
Lys Gly Thr Glu Thr Arg Gly Lys Leu Cys Pro Lys Cys Leu Asn Cys
1 5 10 15
Thr Asp Leu Asp Val Ala Leu Gly Arg Pro Lys Cys Thr Gly Asn Ile
20 25 30
Pro Ser Ala Arg Val Ser Ile Leu His Glu Val Arg Pro Val Thr Ser
35 40 45
Gly Cys Phe Pro Ile Met His Asp Arg Thr Lys Ile Arg Gln Leu Pro
50 55 60
Asn Leu Leu Arg Gly Tyr Glu His Ile Arg Leu Ser Thr His Asn Val
65 70 75 80
Ile Asn Ala Glu Asn Ala Pro Gly Gly Ser Tyr Lys Ile Gly Thr Ser
85 90 95
Gly Ser Cys Pro Asn Val Thr Asn Gly Asn Gly Phe Phe Ala Thr Met
100 105 110
Ala Trp Ala Val Pro Lys Asn Asp Asn Asn Lys Thr Ala Thr Asn Ser
115 120 125
Leu Thr Ile Glu Val Pro Tyr Ile Cys Thr Glu Gly Glu Asp Gln Ile
130 135 140
Thr Val Trp Gly Phe His Ser Asp Asn Glu Ala Gln Met Ala Lys Leu
145 150 155 160
Tyr Gly Asp Ser Lys Pro Gln Lys Phe Thr Ser Ser Ala Asn Gly Val
165 170 175
Thr Thr His Tyr Val Ser Gln Ile Gly Gly Phe Pro Asn Gln Thr Glu
180 185 190
Asp Gly Gly Leu Pro Gln Ser Gly Arg Ile Val Val Asp Tyr Met Val
195 200 205
Gln Lys Ser Gly Lys Thr Gly Thr Ile Thr Tyr Gln Arg Gly Ile Leu
210 215 220
Leu Pro Gln Lys Val Trp Cys Ala Ser Gly Arg Ser Lys Val Ile Lys
225 230 235 240
Gly

<210> 47
<211> 294
<212> PRT
<213> Artificial Sequence

<220>
<223> B/Ohio/1/2005 HA1-1

<400> 47
Thr Thr Thr Pro Thr Lys Ser His Phe Ala Asn Leu Lys Gly Thr Lys
1 5 10 15
Thr Arg Gly Lys Leu Cys Pro Lys Cys Leu Asn Cys Thr Asp Leu Asp
20 25 30
Val Ala Leu Gly Arg Pro Lys Cys Thr Gly Asn Ile Pro Ser Ala Glu

```
        35                  40                  45
Val Ser Ile Leu His Glu Val Arg Pro Val Thr Ser Gly Cys Phe Pro
    50                  55                  60
Ile Met His Asp Arg Thr Lys Ile Arg Gln Leu Pro Asn Leu Leu Arg
65                  70                  75                  80
Gly Tyr Glu His Ile Arg Leu Ser Thr His Asn Val Ile Asn Ala Glu
                85                  90                  95
Lys Ala Pro Gly Gly Pro Tyr Lys Ile Gly Thr Ser Gly Ser Cys Pro
            100                 105                 110
Asn Val Thr Asn Gly Asn Gly Phe Phe Ala Thr Met Ala Trp Ala Val
        115                 120                 125
Pro Lys Asn Asp Asn Asn Lys Thr Ala Thr Asn Ser Leu Thr Ile Glu
    130                 135                 140
Val Pro Tyr Ile Cys Thr Glu Gly Glu Asp Gln Ile Thr Ile Trp Gly
145                 150                 155                 160
Phe His Ser Asp Ser Glu Thr Gln Met Ala Lys Leu Tyr Gly Asp Ser
                165                 170                 175
Lys Pro Gln Lys Phe Thr Ser Ser Ala Asn Gly Val Thr Thr His Tyr
            180                 185                 190
Val Ser Gln Ile Gly Gly Phe Pro Asn Gln Thr Glu Asp Gly Gly Leu
        195                 200                 205
Pro Gln Ser Gly Arg Ile Val Val Asp Tyr Met Val Gln Lys Ser Gly
    210                 215                 220
Lys Thr Gly Thr Ile Thr Tyr Gln Arg Gly Ile Leu Leu Pro Gln Lys
225                 230                 235                 240
Val Trp Cys Ala Ser Gly Arg Ser Lys Val Ile Lys Gly Ser Leu Pro
                245                 250                 255
Leu Ile Gly Glu Ala Asp Cys Leu His Glu Lys Tyr Gly Gly Leu Asn
            260                 265                 270
Lys Ser Lys Pro Tyr Tyr Thr Gly Glu His Ala Lys Ala Ile Gly Asn
        275                 280                 285
Cys Pro Ile Trp Val Lys
    290
```

<210> 48
<211> 241
<212> PRT
<213> Artificial Sequence

<220>
<223> B/Ohio/1/2005 HA1-2

```
<400> 48
Lys Gly Thr Lys Thr Arg Gly Lys Leu Cys Pro Lys Cys Leu Asn Cys
 1               5                  10                  15
Thr Asp Leu Asp Val Ala Leu Gly Arg Pro Lys Cys Thr Gly Asn Ile
            20                  25                  30
Pro Ser Ala Glu Val Ser Ile Leu His Glu Val Arg Pro Val Thr Ser
        35                  40                  45
Gly Cys Phe Pro Ile Met His Asp Arg Thr Lys Ile Arg Gln Leu Pro
    50                  55                  60
Asn Leu Leu Arg Gly Tyr Glu His Ile Arg Leu Ser Thr His Asn Val
65                  70                  75                  80
Ile Asn Ala Glu Lys Ala Pro Gly Gly Pro Tyr Lys Ile Gly Thr Ser
                85                  90                  95
Gly Ser Cys Pro Asn Val Thr Asn Gly Asn Gly Phe Phe Ala Thr Met
            100                 105                 110
Ala Trp Ala Val Pro Lys Asn Asp Asn Asn Lys Thr Ala Thr Asn Ser
```

```
        115                 120                 125
Leu Thr Ile Glu Val Pro Tyr Ile Cys Thr Glu Gly Glu Asp Gln Ile
    130                 135                 140
Thr Ile Trp Gly Phe His Ser Asp Ser Glu Thr Gln Met Ala Lys Leu
145                 150                 155                 160
Tyr Gly Asp Ser Lys Pro Gln Lys Phe Thr Ser Ser Ala Asn Gly Val
                165                 170                 175
Thr Thr His Tyr Val Ser Gln Ile Gly Gly Phe Pro Asn Gln Thr Glu
            180                 185                 190
Asp Gly Gly Leu Pro Gln Ser Gly Arg Ile Val Val Asp Tyr Met Val
        195                 200                 205
Gln Lys Ser Gly Lys Thr Gly Thr Ile Thr Tyr Gln Arg Gly Ile Leu
    210                 215                 220
Leu Pro Gln Lys Val Trp Cys Ala Ser Gly Arg Ser Lys Val Ile Lys
225                 230                 235                 240
Gly
```

<210> 49
<211> 293
<212> PRT
<213> Artificial Sequence

<220>
<223> B/Shanghai/361/2002 HA1-1

```
<400> 49
Thr Thr Thr Pro Ile Lys Ser His Phe Ala Asn Leu Lys Gly Thr Arg
 1               5                  10                  15
Thr Arg Gly Lys Leu Cys Pro Asp Cys Leu Asn Cys Thr Asp Leu Asp
            20                  25                  30
Val Ala Leu Gly Arg Pro Met Cys Val Gly Thr Thr Pro Ser Ala Lys
        35                  40                  45
Ala Ser Ile Leu His Glu Val Arg Pro Val Thr Ser Gly Cys Phe Pro
    50                  55                  60
Ile Met His Asp Arg Thr Lys Ile Arg Gln Leu Pro Asn Leu Leu Arg
65                  70                  75                  80
Gly Tyr Glu Asn Ile Arg Leu Ser Thr Gln Asn Val Ile Asp Ala Glu
                85                  90                  95
Lys Ala Leu Gly Gly Pro Tyr Arg Leu Gly Thr Ser Gly Ser Cys Pro
            100                 105                 110
Asn Ala Thr Ser Lys Ser Gly Phe Phe Ala Thr Met Ala Trp Ala Val
        115                 120                 125
Pro Lys Asp Asn Asn Lys Asn Ala Thr Asn Pro Leu Thr Val Glu Val
    130                 135                 140
Pro Tyr Ile Cys Thr Glu Gly Glu Asp Gln Ile Thr Val Trp Gly Phe
145                 150                 155                 160
His Ser Asp Asp Lys Thr Gln Met Lys Asn Leu Tyr Gly Asp Ser Asn
                165                 170                 175
Pro Gln Lys Phe Thr Ser Ser Ala Asn Gly Val Thr Thr His Tyr Val
            180                 185                 190
Ser Gln Ile Gly Gly Phe Pro Asp Gln Thr Glu Asp Gly Gly Leu Pro
        195                 200                 205
Gln Ser Gly Arg Ile Val Val Asp Tyr Met Val Gln Lys Pro Gly Lys
    210                 215                 220
Thr Gly Thr Ile Val Tyr Gln Arg Gly Val Leu Leu Pro Gln Lys Val
225                 230                 235                 240
Trp Cys Ala Ser Gly Arg Ser Lys Val Ile Lys Gly Ser Leu Pro Leu
```

```
                245                 250                 255
Ile Gly Glu Ala Asp Cys Leu His Glu Lys Tyr Gly Gly Leu Asn Lys
            260                 265                 270
Ser Lys Pro Tyr Tyr Thr Gly Glu His Ala Lys Ala Ile Gly His Cys
        275                 280                 285
Pro Ile Trp Val Lys
    290
```

<210> 50
<211> 240
<212> PRT
<213> Artificial Sequence

<220>
<223> B/Shanghai/361/2002 HA1-2

<400> 50

```
Lys Gly Thr Arg Thr Arg Gly Lys Leu Cys Pro Asp Cys Leu Asn Cys
 1               5                  10                  15
Thr Asp Leu Asp Val Ala Leu Gly Arg Pro Met Cys Val Gly Thr Thr
            20                  25                  30
Pro Ser Ala Lys Ala Ser Ile Leu His Glu Val Arg Pro Val Thr Ser
        35                  40                  45
Gly Cys Phe Pro Ile Met His Asp Arg Thr Lys Ile Arg Gln Leu Pro
    50                  55                  60
Asn Leu Leu Arg Gly Tyr Glu Asn Ile Arg Leu Ser Thr Gln Asn Val
65                  70                  75                  80
Ile Asp Ala Glu Lys Ala Leu Gly Gly Pro Tyr Arg Leu Gly Thr Ser
                85                  90                  95
Gly Ser Cys Pro Asn Ala Thr Ser Lys Ser Gly Phe Phe Ala Thr Met
            100                 105                 110
Ala Trp Ala Val Pro Lys Asp Asn Asn Lys Asn Ala Thr Asn Pro Leu
        115                 120                 125
Thr Val Glu Val Pro Tyr Ile Cys Thr Glu Gly Glu Asp Gln Ile Thr
    130                 135                 140
Val Trp Gly Phe His Ser Asp Asp Lys Thr Gln Met Lys Asn Leu Tyr
145                 150                 155                 160
Gly Asp Ser Asn Pro Gln Lys Phe Thr Ser Ser Ala Asn Gly Val Thr
                165                 170                 175
Thr His Tyr Val Ser Gln Ile Gly Gly Phe Pro Asp Gln Thr Glu Asp
            180                 185                 190
Gly Gly Leu Pro Gln Ser Gly Arg Ile Val Val Asp Tyr Met Val Gln
        195                 200                 205
Lys Pro Gly Lys Thr Gly Thr Ile Val Tyr Gln Arg Gly Val Leu Leu
    210                 215                 220
Pro Gln Lys Val Trp Cys Ala Ser Gly Arg Ser Lys Val Ile Lys Gly
225                 230                 235                 240
```

<210> 51
<211> 1518
<212> DNA
<213> Artificial Sequence

<220>
<223> Stf2.blp

<400> 51

atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcagaa agcgtatgat 540
gtgaaagata cagcagtaac aacgaaagct tatgccaata atggtactac actggatgta 600
tcgggtcttg atgatgcagc tattaaagcg gctacgggtg gtacgaatgg tacggcttct 660
gtaaccggtg gtgcggttaa atttgacgca gataataaca agtactttgt tactattggt 720
ggctttactg gtgctgatgc cgccaaaaat ggcgattatg aagttaacgt tgctactgac 780
ggtacagtaa cccttgcggc tggcgcaact aaaaccacaa tgcctgctgg tgcgacaact 840
aaaacagaag tacaggagtt aaaagataca ccggcagttg tttcagcaga tgctaaaaat 900
gccttaattg ctggcggcgt tgacgctacc gatgctaatg gcgctgagtt ggtcaaaatg 960
tcttataccg ataaaaatgg taagacaatt gaaggcggtt atgcgcttaa agctggcgat 1020
aagtattacg ccgcagatta cgatgaagcg acaggagcaa ttaaagctaa aactacaagt 1080
tatactgctg ctgacggcac taccaaaaca gcggctaacc aactgggtgg cgtagacggt 1140
aaaaccgaag tcgttactat cgacggtaaa acctacaatg ccagcaaagc cgctggtcat 1200
gatttcaaag cacaaccaga gctggcggaa gcagccgcta aaaccaccga aaacccgctg 1260
cagaaaattg atgccgcgct ggcgcaggtg gatgcgctgc gctctgatct gggtgcggta 1320
caaaaccgtt tcaactctgc tatcaccaac ctgggcaata ccgtaaacaa tctgtctgaa 1380
gcgcgtagcc gtatcgaaga ttccgactac gcgaccgaag tttccaacat gtctcgcgcg 1440
cagattctgc agcaggccgg tacttccgtt ctggcgcagg ctaaccaggt cccgcagaac 1500
gtgctgagcc tgttacgt 1518

<210> 52
<211> 1539
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2.SG

<400> 52
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcagaa agcgtatgat 540
gtgaaagata cagcagtaac aacgaaagct tatgccaata atggtactac actggatgta 600
tcgggtcttg atgatgcagc tattaaagcg gctacgggtg gtacgaatgg tacggcttct 660
gtaaccggtg gtgcggttaa atttgacgca gataataaca agtactttgt tactattggt 720
ggctttactg gtgctgatgc cgccaaaaat ggcgattatg aagttaacgt tgctactgac 780
ggtacagtaa cccttgcggc tggcgcaact aaaaccacaa tgcctgctgg tgcgacaact 840
aaaacagaag tacaggagtt aaaagataca ccggcagttg tttcagcaga tgctaaaaat 900
gccttaattg ctggcggcgt tgacgctacc gatgctaatg gcgctgagtt ggtcaaaatg 960
tcttataccg ataaaaatgg taagacaatt gaaggcggtt atgcgcttaa agctggcgat 1020
aagtattacg ccgcagatta cgatgaagcg acaggagcaa ttaaagctaa aactacaagt 1080
tatactgctg ctgacggcac taccaaaaca gcggctaacc aactgggtgg cgtagacggt 1140
aaaaccgaag tcgttactat cgacggtaaa acctacaatg ccagcaaagc cgctggtcat 1200
gatttcaaag cacaaccaga gctggcggaa gcagccgcta aaaccaccga aaacccgctg 1260
cagaaaattg atgccgcgct ggcgcaggtg gatgcgctgc gctctgatct gggtgcggta 1320
caaaaccgtt tcaactctgc tatcaccaac ctgggcaata ccgtaaacaa tctgtctgaa 1380

gcgcgtagcc gtatcgaaga ttccgactac gcgaccgaag tttccaacat gtctcgcgcg 1440
cagattctgc agcaggccgg tacttccgtt ctggcgcagg ctaaccaggt cccgcagaac 1500
gtgctgtctc tgttacgtag cggcagtggt agcggatcc 1539

<210> 53
<211> 2361
<212> DNA
<213> Artificial Sequence

<220>
<223> SFT2.HA1-1 PR8

<400> 53
atggcacaag taatcaacac taacagtctg tcgctgctga.cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcagaa agcgtatgat 540
gtgaaagata cagcagtaac aacgaaagct tatgccaata atggtactac actggatgta 600
tcgggtcttg atgatgcagc tattaaagcg gctacgggtg gtacgaatgg tacggcttct 660
gtaaccggtg gtgcggttaa atttgacgca gataataaca agtactttgt tactattggt 720
ggctttactg gtgctgatgc cgccaaaaat ggcgattatg aagttaacgt tgctactgac 780
ggtacagtaa cccttgcggc tggcgcaact aaaaccacaa tgcctgctgg tgcgacaact 840
aaaacagaag tacaggagtt aaaagataca ccggcagttg tttcagcaga tgctaaaaat 900
gccttaattg ctggcggcgt tgacgctacc gatgctaatg gcgctgagtt ggtcaaaatg 960
tcttataccg ataaaaatgg taagacaatt gaaggcggtt atgcgcttaa agctggcgat 1020
aagtattacg ccgcagatta cgatgaagcg acaggagcaa ttaaagctaa aaccacaagt 1080
tatactgctg ctgacggcac taccaaaaca gcggctaacc aactgggtgg cgtagacggt 1140
aaaaccgaag tcgttactat cgacggtaaa acctacaatg ccagcaaagc cgctggtcat 1200
gatttcaaag cacaaccaga gctggcggaa gcagccgcta aaaccaccga aaacccgctg 1260
cagaaaattg atgccgcgct ggcgcaggtg gatgcgctgc gctctgatct gggtgcggta 1320
caaaaccgtt tcaactctgc tatcaccaac ctgggcaata ccgtaaacaa tctgtctgaa 1380
gcgcgtagcc gtatcgaaga ttccgactac gcgaccgaag tttccaacat gtctcgcgcg 1440
cagattttgc agcaggccgg tacttccgtt ctggcgcagg ctaaccaggt cccgcagaac 1500
gtgctgtctc tgttacgtag cggcagtggt agcggatcct ctcataacgg taaactgtgt 1560
cgtctgaaag gtattgcacc actgcagctg ggcaaatgca acatcgcggg ttggctgctg 1620
ggtaaccctg aatgtgaccc gctgctgccg gttcgttcct ggagctacat tgttgaaacc 1680
ccgaactccg aaaacggtat ctgctacccg ggcgacttta ttgactatga agaactgcgt 1740
gagcagctgt cttccgtgag cagctttgaa cgcttcgaaa tcttcccgaa ggaaagctcc 1800
tggccgaacc acaacactaa cggcgtgacg gcggcttgct cccacgaagg caaatcttcc 1860
ttttatcgta acctgctgtg gctgactgaa aaggaaggtt cctacccaaa actgaaaaac 1920
agctatgtta acaaaaaggg taaagaagtc ctggtgctgt ggggcatcca ccacccgccg 1980
aactccaagg aacagcagaa tctgtatcag aacgaaaacg catacgtttc tgtcgttact 2040
tccaactata accgtcgttt cactccggaa atcgcggaac gtccgaaagt acgcgaccag 2100
gctggccgta tgaactacta ctggaccctg ctgaaaccgg gtgacactat tatcttcgaa 2160
gctaacggta acctgatcgc accaatgtac gctttcgcac tgtctcgtgg tttcggttcc 2220
ggcattatca ccagcaacgc gtctatgcac gaatgcaaca cgaaatgtca gacgccgctg 2280
ggcgcaatta atagcagcct gccgtaccag aacatccacc cggtgactat cggcgaatgc 2340
ccgaagtatg ttcgttaata g 2361

<210> 54
<211> 2379
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2.HA1-1.his PR8

<400> 54
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcagaa agcgtatgat 540
gtgaaagata cagcagtaac aacgaaagct tatgccaata atggtactac actggatgta 600
tcgggtcttg atgatgcagc tattaaagcg gctacgggtg gtacgaatgg tacggcttct 660
gtaaccggtg gtgcggttaa atttgacgca gataataaca agtactttgt tactattggt 720
ggctttactg gtgctgatgc cgccaaaaat ggcgattatg aagttaacgt tgctactgac 780
ggtacagtaa cccttgcggc tggcgcaact aaaaccacaa tgcctgctgg tgcgacaact 840
aaaacagaag tacaggagtt aaaagataca ccggcagttg tttcagcaga tgctaaaaat 900
gccttaattg ctggcggcgt tgacgctacc gatgctaatg gcgctgagtt ggtcaaaatg 960
tcttataccg ataaaaatgg taagacaatt gaaggcggtt atgcgcttaa agctggcgat 1020
aagtattacg ccgcagatta cgatgaagcg acaggagcaa ttaaagctaa aaccacaagt 1080
tatactgctg ctgacggcac taccaaaaca gcggctaacc aactgggtgg cgtagacggt 1140
aaaaccgaag tcgttactat cgacggtaaa acctacaatg ccagcaaagc cgctggtcat 1200
gatttcaaag cacaaccaga gctggcggaa gcagccgcta aaaccaccga aaacccgctg 1260
caaaaaattg atgccgcgct ggcgcaggtg gatgcgctgc gctctgatct gggtgcggta 1320
caaaaccgtt tcaactctgc tatcaccaac ctgggcaata ccgtaaacaa tctgtctgaa 1380
gcgcgtagcc gtatcgaaga ttccgactac gcgaccgaag tttccaacat gtctcgcgcg 1440
cagattttgc agcaggccgg tacttccgtt ctggcgcagg ctaaccaggt cccgcagaac 1500
gtgctgtctc tgttacgtag cggcagtggt agcggatcct ctcataacgg taaactgtgt 1560
cgtctgaaag gtattgcacc actgcagctg ggcaaatgca acatcgcggg ttggctgctg 1620
ggtaaccctg aatgtgaccc gctgctgccg gttcgttcct ggagctacat tgttgaaacc 1680
ccgaactccg aaaacggtat ctgctacccg ggcgacttta ttgactatga agaactgcgt 1740
gagcagctgt cttccgtgag cagctttgaa cgcttcgaaa tcttcccgaa ggaaagctcc 1800
tggccgaacc acaacactaa cggcgtgacg gcggcttgct cccacgaagg caaatcttcc 1860
ttttatcgta acctgctgtg gctgactgaa aaggaaggtt cctacccaaa actgaaaaac 1920
agctatgtta acaaaaaggg taaagaagtc ctggtgctgt ggggcatcca ccacccgccg 1980
aactccaagg aacagcagaa tctgtatcag aacgaaaacg catacgtttc tgtcgttact 2040
tccaactata accgtcgttt cactccggaa atcgcggaac gtccgaaagt acgcgaccag 2100
gctggccgta tgaactacta ctggaccctg ctgaaaccgg gtgacactat tatcttcgaa 2160
gctaacggta acctgatcgc accaatgtac gctttcgcac tgtctcgtgg tttcggttcc 2220
ggcattatca ccagcaacgc gtctatgcac gaatgcaaca cgaaatgtca gacgccgctg 2280
ggcgcaatta atagcagcct gccgtaccag aacatccacc cggtgactat cggcgaatgc 2340
ccgaagtatg ttcgtcatca ccatcatcac cattaatag 2379

<210> 55
<211> 2214
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2.HA1-2 PR8

<400> 55
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300

aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcagaa agcgtatgat 540
gtgaaagata cagcagtaac aacgaaagct tatgccaata atggtactac actggatgta 600
tcgggtcttg atgatgcagc tattaaagcg gctacgggtg gtacgaatgg tacggcttct 660
gtaaccggtg gtgcggttaa atttgacgca gataataaca agtactttgt tactattggt 720
ggctttactg gtgctgatgc cgccaaaaat ggcgattatg aagttaacgt tgctactgac 780
ggtacagtaa cccttgcggc tggcgcaact aaaaccacaa tgcctgctgg tgcgacaact 840
aaaacagaag tacaggagtt aaaagataca ccggcagttg tttcagcaga tgctaaaaat 900
gccttaattg ctggcggcgt tgacgctacc gatgctaatg gcgctgagtt ggtcaaaatg 960
tcttataccg ataaaaatgg taagacaatt gaaggcggtt atgcgcttaa agctggcgat 1020
aagtattacg ccgcagatta cgatgaagcg acaggagcaa ttaaagctaa aactacaagt 1080
tatactgctg ctgacggcac taccaaaaca gcggctaacc aactgggtgg cgtagacggt 1140
aaaaccgaag tcgttactat cgacggtaaa acctacaatg ccagcaaagc cgctggtcat 1200
gatttcaaag cacaaccaga gctggcggaa gcagccgcta aaaccaccga aaacccgctg 1260
cagaaaattg atgccgcgct ggcgcaggtg gatgcgctgc gctctgatct gggtgcggta 1320
caaaaccgtt tcaactctgc tatcaccaac ctgggcaata ccgtaaacaa tctgtctgaa 1380
gcgcgtagcc gtatcgaaga ttccgactac gcgaccgaag tttccaacat gtctcgcgcg 1440
cagattctgc agcaggccgg tacttccgtt ctggcgcagg ctaaccaggt cccgcagaac 1500
gtgctgtctc tgttacgttc tggttctggt tctggttcta aaggtattgc tccactgcaa 1560
ctgggtaaat gcaatattgc gggttggctg ctgggcaacc cggaatgcga tccgctgctg 1620
ccggtccgtt cctggagcta tattgttgaa actccgaact ctgagaacgg catctgctat 1680
ccaggtgatt tcattgacta tgaggaactg cgtgaacaac tgtcttccgt gtcttccttt 1740
gaacgtttcg agatttttcc taaagaatct tcttggccga accataacac taatggtgtt 1800
accgctgcgt gctctcatga aggtaaatct agcttttacc gcaacctgct gtggctgacc 1860
gagaaagaag gttcttaccc gaaactgaaa aacagctacg taaacaaaaa gggcaaggaa 1920
gttctggtcc tgtggggtat ccaccatccg ccgaacagca aggaacagca gaatctgtat 1980
cagaacgaaa acgcatacgt atctgttgtt acttctaact acaaccgccg tttcacccct 2040
gaaatcgcgg aacgtccgaa agtgcgtgac caggcaggcc gcatgaacta ttactggacc 2100
ctgctgaagc cgggtgatac tatcatcttc gaagcgaacg gtaacctgat cgccccgatg 2160
tacgcgttcg ctctgagccg tggcttcggc tctggtatca ttacgtctta ataa 2214

<210> 56
<211> 2214
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2.HA1-2mut PR8

<400> 56
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcagaa agcgtatgat 540
gtgaaagata cagcagtaac aacgaaagct tatgccaata atggtactac actggatgta 600
tcgggtcttg atgatgcagc tattaaagcg gctacgggtg gtacgaatgg tacggcttct 660
gtaaccggtg gtgcggttaa atttgacgca gataataaca agtactttgt tactattggt 720
ggctttactg gtgctgatgc cgccaaaaat ggcgattatg aagttaacgt tgctactgac 780
ggtacagtaa cccttgcggc tggcgcaact aaaaccacaa tgcctgctgg tgcgacaact 840
aaaacagaag tacaggagtt aaaagataca ccggcagttg tttcagcaga tgctaaaaat 900
gccttaattg ctggcggcgt tgacgctacc gatgctaatg gcgctgagtt ggtcaaaatg 960
tcttataccg ataaaaatgg taagacaatt gaaggcggtt atgcgcttaa agctggcgat 1020

aagtattacg ccgcagatta cgatgaagcg acaggagcaa ttaaagctaa aaccacaagt 1080
tatactgctg ctgacggcac taccaaaaca gcggctaacc aactgggtgg cgtagacggt 1140
aaaaccgaag tcgttactat cgacggtaaa acctacaatg ccagcaaagc cgctggtcat 1200
gatttcaaag cacaaccaga gctggcggaa gcagccgcta aaaccaccga aaacccgctg 1260
cagaaaattg atgccgcgct ggcgcaggtg gatgcgctgc gctctgatct gggtgcggta 1320
caaaaccgtt tcaactctgc tatcaccaac ctgggcaata ccgtaaacaa tctgtctgaa 1380
gcgcgtagcc gtatcgaaga ttccgactac gcgaccgaag tttccaacat gtctcgcgcg 1440
cagattttgc agcaggccgg tacttccgtt ctggcgcagg ctaaccaggt cccgcagaac 1500
gtgctgtctc tgttagcgag cggcagtggt agcggatcca aaggtgcggc tccactgcaa 1560
ctgggtaaat gcaatattgc gggttggctg ctgggcaacc cggaatgcga tccgctgctg 1620
ccggtccgtt cctggagcga tattgcggaa actccgaact ctgagaacgg catctgctat 1680
ccaggtgatt tcattgacta tgaggaactg cgtgaacaac tgtcttccgt gtcttccttt 1740
gaacgtttcg agatttttcc taaagaatct tcttggccga accataacac taatggtgtt 1800
accgctgcgt gctctcatga aggtaaatct agcttttacc gcaacctgct gtggctgacc 1860
gagaaagaag gttcttaccc gaaactgaaa aacagctacg taaacaaaaa gggcaaggaa 1920
gttctggtcc tgtggggtat ccaccatccg ccgaacagca aggaacagca gaatctgtat 1980
cagaacgaaa acgcatacgt atctgttgtt acttctaact acaaccgccg tttcacccct 2040
gaaatcgcgg aacgtccgaa agtgcgtgac caggcaggcc gcatgaacta ttactggacc 2100
ctgctgaagc cgggtgatac tatcatcttc gaagcgaacg gtaacctgat cgccccgatg 2160
tacgcgttcg ctctgagccg tggcttcggc tctggtatca ttacgtctta ataa 2214

<210> 57
<211> 2073
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2.HA1-3 PR8

<400> 57
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcagaa agcgtatgat 540
gtgaaagata cagcagtaac aacgaaagct tatgccaata atggtactac actggatgta 600
tcgggtcttg atgatgcagc tattaaagcg gctacgggtg gtacgaatgg tacggcttct 660
gtaaccggtg gtgcggttaa atttgacgca gataataaca agtactttgt tactattggt 720
ggctttactg gtgctgatgc cgccaaaaat ggcgattatg aagttaacgt tgctactgac 780
ggtacagtaa cccttgcggc tggcgcaact aaaaccacaa tgcctgctgg tgcgacaact 840
aaaacagaag tacaggagtt aaaagataca ccggcagttg tttcagcaga tgctaaaaat 900
gccttaattg ctggcggcgt tgacgctacc gatgctaatg gcgctgagtt ggtcaaaatg 960
tcttataccg ataaaaatgg taagacaatt gaaggcggtt atgcgcttaa agctggcgat 1020
aagtattacg ccgcagatta cgatgaagcg acaggagcaa ttaaagctaa aactacaagt 1080
tatactgctg ctgacggcac taccaaaaca gcggctaacc aactgggtgg cgtagacggt 1140
aaaaccgaag tcgttactat cgacggtaaa acctacaatg ccagcaaagc cgctggtcat 1200
gatttcaaag cacaaccaga gctggcggaa gcagccgcta aaaccaccga aaacccgctg 1260
cagaaaattg atgccgcgct ggcgcaggtg gatgcgctgc gctctgatct gggtgcggta 1320
caaaaccgtt tcaactctgc tatcaccaac ctgggcaata ccgtaaacaa tctgtctgaa 1380
gcgcgtagcc gtatcgaaga ttccgactac gcgaccgaag tttccaacat gtctcgcgcg 1440
cagattctgc agcaggccgg tacttccgtt ctggcgcagg ctaaccaggt cccgcagaac 1500
gtgctgtctc tgttacgttc tggtagcggt tctggctcta actctgaaaa tggtatctgt 1560
tacccgggtg atttcatcga ttatgaggaa ctgcgtgaac agctgtctag cgtgagctcc 1620
tttgaacgct tcgaaatctt cccgaaggaa tctagctggc cgaaccataa cacgaacggt 1680
gtgaccgctg cttgctccca cgaaggtaag agctccttct accgcaatct gctgtggctg 1740

actgaaaaag aaggtagcta cccgaaactg aaaaattctt acgtcaacaa gaaaggcaag 1800
gaagtgctgg ttctgtgggg cattcaccac ccaccgaaca gcaaagagca acagaacctg 1860
taccaaaatg agaacgctta cgtttctgtt gtgacttcta actacaatcg tcgctttacc 1920
cctgaaatcg cggagcgtcc aaaagtgcgt gaccaggctg gtcgtatgaa ctactattgg 1980
accctgctga aaccgggcga caccattatc tttgaagcga acggtaacct gatcgcgcct 2040
atgtacgcgt tcgctctgtc tcgtggctaa taa 2073

<210> 58
<211> 2073
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2.HA1-3mut PR8

<400> 58
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcagaa agcgtatgat 540
gtgaaagata cagcagtaac aacgaaagct tatgccaata atggtactac actggatgta 600
tcgggtcttg atgatgcagc tattaaagcg gctacgggtg gtacgaatgg tacggcttct 660
gtaaccggtg gtgcggttaa atttgacgca gataataaca agtactttgt tactattggt 720
ggctttactg gtgctgatgc cgccaaaaat ggcgattatg aagttaacgt tgctactgac 780
ggtacagtaa cccttgcggc tggcgcaact aaaaccacaa tgcctgctgg tgcgacaact 840
aaaacagaag tacaggagtt aaaagataca ccggcagttg tttcagcaga tgctaaaaat 900
gccttaattg ctggcggcgt tgacgctacc gatgctaatg gcgctgagtt ggtcaaaatg 960
tcttataccg ataaaaatgg taagacaatt gaaggcggtt atgcgcttaa agctggcgat 1020
aagtattacg ccgcagatta cgatgaagcg acaggagcaa ttaaagctaa aactacaagt 1080
tatactgctg ctgacggcac taccaaaaca gcggctaacc aactgggtgg cgtagacggt 1140
aaaaccgaag tcgttactat cgacggtaaa acctacaatg ccagcaaagc cgctggtcat 1200
gatttcaaag cacaaccaga gctggcggaa gcagccgcta aaaccaccga aaacccgctg 1260
cagaaaattg atgccgcgct ggcgcaggtg gatgcgctgc gctctgatct gggtgcggta 1320
caaaaccgtt tcaactctgc tatcaccaac ctgggcaata ccgtaaacaa tctgtctgaa 1380
gcgcgtagcc gtatcgaaga ttccgactac gcgaccgaag tttccaacat gtctcgcgcg 1440
cagattctgc agcaggccgg tacttccgtt ctggcgcagg ctaaccaggt cccgcagaac 1500
gtgctgtctc tgttacgtag cggcagtggt agcggatcca actctgaaaa tgaaatctgt 1560
tacccgggtg atttcatcga taaagaggaa ctgcgtgaac agctgtctag cgtgagctcc 1620
tttgaacgct tcgaaatctt cccgaaggaa tctagctggc cgaaccataa cacgaacggt 1680
gtgaccgctg cttgctccca cgaaggtaag agctccttct accgcaatct gctgtggctg 1740
actgaaaaag aaggtagcta cccgaaactg aaaaattctt acgtcaacaa gaaaggcaag 1800
gaagtgctgg ttctgtgggg cattcaccac ccaccgaaca gcaaagagca acagaacctg 1860
taccaaaatg agaacgctta cgtttctgtt gtgacttcta actacaatcg tcgctttacc 1920
cctgaaatcg cggagcgtcc aaaagtgcgt gaccaggctg gtcgtatgaa ctactattgg 1980
accctgctga aaccgggcga caccattatc tttgaagcga acggtaacct gatcgcgcct 2040
atgtacgcgg cggctctgtc tcgtggctaa taa 2073

<210> 59
<211> 819
<212> DNA
<213> Artificial Sequence

<220>
<223> HA1-1 PR8

<400> 59
atgtctcata acggtaaact gtgtcgtctg aaaggtattg caccactgca gctgggcaaa 60
tgcaacatcg cgggttggct gctgggtaac cctgaatgtg acccgctgct gccggttcgt 120
tcctggagct acattgttga aaccccgaac tccgaaaacg gtatctgcta cccgggcgac 180
tttattgact atgaagaact gcgtgagcag ctgtcttccg tgagcagctt tgaacgcttc 240
gaaatcttcc cgaaggaaag ctcctggccg aaccacaaca ctaacggcgt gacggcggct 300
tgctcccacg aaggcaaatc ttccttttat cgtaacctgc tgtggctgac tgaaaaggaa 360
ggttcctacc caaaactgaa aaacagctat gttaacaaaa agggtaaaga agtcctggtg 420
ctgtggggca tccaccaccc gccgaactcc aaggaacagc agaatctgta tcagaacgaa 480
aacgcatacg tttctgtcgt tacttccaac tataaccgtc gtttcactcc ggaaatcgcg 540
gaacgtccga aagtacgcga ccaggctggc cgtatgaact actactggac cctgctgaaa 600
ccgggtgaca ctattatctt cgaagctaac ggtaacctga tcgcaccaat gtacgctttc 660
gcactgtctc gtggtttcgg ttccggcatt atcaccagca acgcgtctat gcacgaatgc 720
aacacgaaat gtcagacgcc gctgggcgca attaatagca gcctgccgta ccagaacatc 780
cacccggtga ctatcggcga atgcccgaag tatgttcgt 819

<210> 60
<211> 837
<212> DNA
<213> Artificial Sequence

<220>
<223> HA1-1.his PR8

<400> 60
atgtctcata acggtaaact gtgtcgtctg aaaggtattg caccactgca gctgggcaaa 60
tgcaacatcg cgggttggct gctgggtaac cctgaatgtg acccgctgct gccggttcgt 120
tcctggagct acattgttga aaccccgaac tccgaaaacg gtatctgcta cccgggcgac 180
tttattgact atgaagaact gcgtgagcag ctgtcttccg tgagcagctt tgaacgcttc 240
gaaatcttcc cgaaggaaag ctcctggccg aaccacaaca ctaacggcgt gacggcggct 300
tgctcccacg aaggcaaatc ttccttttat cgtaacctgc tgtggctgac tgaaaaggaa 360
ggttcctacc caaaactgaa aaacagctat gttaacaaaa agggtaaaga agtcctggtg 420
ctgtggggca tccaccaccc gccgaactcc aaggaacagc agaatctgta tcagaacgaa 480
aacgcatacg tttctgtcgt tacttccaac tataaccgtc gtttcactcc ggaaatcgcg 540
gaacgtccga aagtacgcga ccaggctggc cgtatgaact actactggac cctgctgaaa 600
ccgggtgaca ctattatctt cgaagctaac ggtaacctga tcgcaccaat gtacgctttc 660
gcactgtctc gtggtttcgg ttccggcatt atcaccagca acgcgtctat gcacgaatgc 720
aacacgaaat gtcagacgcc gctgggcgca attaatagca gcctgccgta ccagaacatc 780
cacccggtga ctatcggcga atgcccgaag tatgttcgtc atcaccatca tcaccat 837

<210> 61
<211> 690
<212> DNA
<213> Artificial Sequence

<220>
<223> HA1-2.his PR8

<400> 61
atgaaaggta ttgctccact gcaactgggt aaatgcaata ttgcgggttg gctgctgggc 60
aacccggaat gcgatccgct gctgccggtc cgttcctgga gctatattgt tgaaactccg 120
aactctgaga acggcatctg ctatccaggt gatttcattg actatgagga actgcgtgaa 180
caactgtctt ccgtgtcttc ctttgaacgt ttcgagattt ttcctaaaga atcttcttgg 240
ccgaaccata acactaatgg tgttaccgct gcgtgctctc atgaaggtaa atctagcttt 300
taccgcaacc tgctgtggct gaccgagaaa gaaggttctt acccgaaact gaaaaacagc 360
tacgtaaaca aaaagggcaa ggaagttctg gtcctgtggg gtatccacca tccgccgaac 420
agcaaggaac agcagaatct gtatcagaac gaaaacgcat acgtatctgt tgttacttct 480
aactacaacc gccgtttcac ccctgaaatc gcggaacgtc cgaaagtgcg tgaccaggca 540

ggccgcatga actattactg gaccctgctg aagccgggtg atactatcat cttcgaagcg 600
aacggtaacc tgatcgcccc gatgtacgcg ttcgctctga gccgtggctt cggctctggt 660
atcattacgt ctcatcacca tcatcaccat 690

<210> 62
<211> 2304
<212> DNA
<213> Artificial Sequence

<220>
<223> CRM.HA1-2

<400> 62
atgggtgctg atgatgtcgt tgattcctcc aaaagcttcg ttatggaaaa tttctcttct 60
tatcacggca ccaaaccggg ttatgtggat tctatccaaa aaggcatcca gaaaccgaag 120
tccggtacgc agggtaacta tgatgacgat tggaaaggtt tttactccac cgataataaa 180
tatgacgcgg ctggctactc tgttgacaac gaaaatccac tgtctggtaa ggctggcggt 240
gtcgtaaagg taacgtatcc tggcctgacc aaggtgctgg cactgaaggt ggataacgct 300
gaaaccatca aaaaggagct gggcctgagc ctgactgaac cgctgatgga gcaggtcggt 360
accgaggaat tcatcaaacg cttcggtgat ggcgcctccc gtgttgtgct gtccctgccg 420
ttcgcagaag gctcttcctc tgtcgaatat atcaacaact gggaacaggc taaagctctg 480
agcgtcgaac tggaaattaa ctttgagacc cgtggcaagc gtggtcagga cgcgatgtac 540
gaatatatgg ctcaggcttg tgccggtaac cgtgttcgcc gttccgtcgg ttcctctctg 600
tcttgcatca acctggattg ggacgttatc cgtgataaga ccaaaaccaa aattgaaagc 660
ctgaaggaac acggtccgat caaaaacaaa atgtctgaat ctccgaacaa aaccgtgtcc 720
gaggaaaaag cgaaacagta tctggaagaa ttccaccaga ctgccctgga acatcctgaa 780
ctgtccgaac tgaagactgt aaccggcact aacccggtgt tcgcaggcgc aaactacgcc 840
gcgtgggcgg taaacgttgc gcaggttatt gatagcgaaa ccgcagataa cctggaaaaa 900
acgaccgcag ctctgtctat cctgccgggt atcggttccg ttatgggtat tgcggacggc 960
gctgtgcacc acaacacgga agaaatcgtc gcacagtcta tcgcgctgtc ttctctgatg 1020
gttgctcagg caattccact ggtaggtgaa ctggtggaca ttggctttgc ggcgtacaac 1080
ttcgtcgaaa gcattatcaa cctgttccag gttgtacaca actcttacaa ccgtccggcc 1140
tacagccctg gccacaaaac ccaaccgttt ctgcacgacg gttatgcggt gtcctggaac 1200
acggttgaag actctatcat tcgtaccggc tttcagggcg agtccggcca cgatatcaaa 1260
attactgcag aaaacactcc gctgccgatc gctggcgttc tgctgccgac catcccgggt 1320
aagctggatg tgaacaaaag caaaacccac atctctgtta acggtcgtaa aattcgcatg 1380
cgctgtcgtg ctatcgacgg tgatgttacc ttctgccgtc cgaaatctcc agtctacgtg 1440
ggcaacggtg ttcatgccaa cctgcacgtg gcgttccatc gtagctctag cgaaaaaatc 1500
cactccaacg aaatctctag cgattctatc ggtgttctgg gttatcagaa aacggtggat 1560
catacgaaag tcaattctaa actgagcctg ttcttcgaaa tcaaatctag cggctctgga 1620
tccggttcca aaggcatcgc gccgctgcag ctgggtaaat gtaacattgc gggctggctg 1680
ctgggtaatc cggaatgcga tccgctgctg ccggtccgta gctggtctta cattgttgaa 1740
actccgaact ctgagaatgg catctgctac ccgggcgatt ttatcgacta tgaagaactg 1800
cgtgaacagc tgtcttccgt ttcttccttt gaacgtttcg aaatcttccc gaaagaaagc 1860
agctggccga atcacaatac gaacggtgtt actgctgcgt gttctcatga aggtaaatcc 1920
agcttctacc gtaacctgct gtggctgacc gaaaaagagg gttcttatcc taaactgaaa 1980
aacagctacg ttaacaaaaa gggcaaagaa gtgctggtgc tgtggggtat ccatcaccct 2040
ccgaactcta aagaacaaca gaatctgtat cagaacgaaa acgcttacgt ttccgtggtg 2100
acctctaact ataaccgtcg ttttaccccg gagattgctg aacgtccgaa agtgcgcgat 2160
caggctggcc gtatgaacta ctattggacc ctgctgaaac cgggcgatac catcattttc 2220
gaagctaacg gcaacctgat tgctccgatg tatgcgtttg ctctgtctcg tggcttcggc 2280
tctggtatta ttacgtctta ataa 2304

<210> 63
<211> 1068
<212> DNA
<213> Artificial Sequence

<220>

<223> LTB.HA1-2

<400> 63
atgaataagg ttaagttcta cgtactgttt accgcgctgc tgtcttctct gtgcgcgcat 60
ggtgctccgc agtctattac tgaactgtgc tctgaatacc acaacaccca gatctatact 120
atcaacgata aaatcctgag ctataccgaa tctatggcag gcaaacgcga aatggttatc 180
attaccttta aaagcggcgc cacctttcaa gtggaagttc cgggctctca gcatattgac 240
tctcaaaaaa aagcgatcga acgtatgaaa gatactctgc gcattaccta cctgaccgaa 300
accaaaatcg ataaactgtg cgtatggaac aataagactc ctaactctat cgcagctatt 360
tctatggaaa actctggtag cggatccggt tctaaaggca tcgcgccgct gcagctgggt 420
aaatgtaaca ttgcgggctg gctgctgggt aatccggaat gcgatccgct gctgccggtc 480
cgtagctggt cttacattgt tgaaactccg aactctgaga atggcatctg ctacccgggc 540
gattttatcg actatgaaga actgcgtgaa cagctgtctt ccgtttcttc ctttgaacgt 600
ttcgaaatct tcccgaaaga aagcagctgg ccgaatcaca atacgaacgg tgttactgct 660
gcgtgttctc atgaaggtaa atccagcttc taccgtaacc tgctgtggct gaccgaaaaa 720
gagggttctt atcctaaact gaaaaacagc tacgttaaca aaaagggcaa agaagtgctg 780
gtgctgtggg gtatccatca ccctccgaac tctaaagaac aacagaatct gtatcagaac 840
gaaaacgctt acgtttccgt ggtgacctct aactataacc gtcgttttac cccggagatt 900
gctgaacgtc cgaaagtgcg cgatcaggct ggccgtatga actactattg gaccctgctg 960
aaaccgggcg ataccatcat tttcgaagct aacggcaacc tgattgctcc gatgtatgcg 1020
tttgctctgt ctcgtggctt cggctctggt attattacgt cttaataa 1068

<210> 64
<211> 2346
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2.HA1-1 VN

<400> 64
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcagaa agcgtatgat 540
gtgaaagata cagcagtaac aacgaaagct tatgccaata atggtactac actggatgta 600
tcgggtcttg atgatgcagc tattaaagcg gctacgggtg gtacgaatgg tacggcttct 660
gtaaccggtg gtgcggttaa atttgacgca gataataaca agtactttgt tactattggt 720
ggctttactg gtgctgatgc cgccaaaaat ggcgattatg aagttaacgt tgctactgac 780
ggtacagtaa cccttgcggc tggcgcaact aaaaccacaa tgcctgctgg tgcgacaact 840
aaaacagaag tacaggagtt aaaagataca ccggcagttg tttcagcaga tgctaaaaat 900
gccttaattg ctggcggcgt tgacgctacc gatgctaatg gcgctgagtt ggtcaaaatg 960
tcttataccg ataaaaatgg taagacaatt gaaggcggtt atgcgcttaa agctggcgat 1020
aagtattacg ccgcagatta cgatgaagcg acaggagcaa ttaaagctaa aaccacaagt 1080
tatactgctg ctgacggcac taccaaaaca gcggctaacc aactgggtgg cgtagacggt 1140
aaaaccgaag tcgttactat cgacggtaaa acctacaatg ccagcaaagc cgctggtcat 1200
gatttcaaag cacaaccaga gctggcggaa gcagccgcta aaaccaccga aaacccgctg 1260
cagaaaattg atgccgcgct ggcgcaggtg gatgcgctgc gctctgatct gggtgcggta 1320
caaaaccgtt tcaactctgc tatcaccaac ctgggcaata ccgtaaacaa tctgtctgaa 1380
gcgcgtagcc gtatcgaaga ttccgactac gcgaccgaag tttccaacat gtctcgcgcg 1440
cagattttgc agcaggccgg tacttccgtt ctggcgcagg ctaaccaggt cccgcagaac 1500
gtgctgagcc tgttagcgga gaagaaacac aatggcaaac tgtgtgatct ggatggtgtg 1560
aaaccgctga ttctgcgcga ttgctctgtg gcaggctggc tgctgggcaa ccctatgtgt 1620
gacgaattca ttaacgttcc ggaatggtct tacattgttg aaaaagctaa ccctgtcaac 1680

gatctgtgtt accctggtga ctttaacgat tacgaagaac tgaagcacct gctgtctcgt 1740
atcaatcact tcgagaaaat ccagatcatc ccgaaatcct cctggagctc ccacgaagct 1800
tctctgggcg tatcctccgc gtgcccgtac cagggcaaat cctctttctt tcgtaacgtt 1860
gtttggctga tcaagaaaaa ctccacctac ccgacgatca agcgtagcta taataacacc 1920
aaccaggaag acctgctggt tctgtggggc atccaccatc caaacgatgc tgcggaacag 1980
accaagctgt accagaaccc gaccacctac atcagcgtgg gcacctctac gctgaaccag 2040
cgtctggtac cgcgtatcgc aacccgcagc aaggtaaacg gtcaaagcgg ccgcatggaa 2100
tttttctgga ccatcctgaa accgaacgac gcaatcaact tcgaatctaa cggcaatttc 2160
atcgctccgg agtatgcgta caaaatcgta aagaaaggtg atagcactat catgaaatcc 2220
gagctggaat atggcaactg taacaccaaa tgccagaccc cgatgggtgc aatcaactcc 2280
tccatgccgt ttcacaacat tcacccgctg actatcggcg aatgtccgaa atacgttaaa 2340
tagtaa 2346

<210> 65
<211> 46
<212> DNA
<213> Artificial Sequence

<220>
<223> FOR PRIMER

<400> 65
agtggctgag cctgttagcg gagaagaaac acaatggcaa actgtg 46

<210> 66
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> REV PRIMER

<400> 66
agtcgctcag cttactattt aacgtatttc ggacattcgc cgatagtcag 50

<210> 67
<211> 1605
<212> DNA
<213> Artificial Sequence

<220>
<223> HA0s VN

<400> 67
gtgctgagcc tgttacgtca gatttgtatc ggctaccacg caaacaactc taccgagcaa 60
gttgatacca tcatggagaa aaacgtgacc gttactcacg cgcaggacat cctggagaag 120
aaacacaatg gcaaactgtg tgatctggat ggtgtgaaac cgctgattct gcgcgattgc 180
tctgtggcag gctggctgct gggcaaccct atgtgtgacg aattcattaa cgttccggaa 240
tggtcttaca ttgttgaaaa agctaaccct gtcaacgatc tgtgttaccc tggtgacttt 300
aacgattacg aagaactgaa gcacctgctg tctcgtatca atcacttcga gaaaatccag 360
atcatcccga aatcctcctg gagctcccac gaagcttctc tgggcgtatc ctccgcgtgc 420
ccgtaccagg gcaaatcctc tttctttcgt aacgttgttt ggctgatcaa gaaaaactcc 480
acctacccga cgatcaagcg tagctataat aacaccaacc aggaagacct gctggttctg 540
tggggcatcc accatccaaa cgatgctgcg gaacagacca agctgtacca gaacccgacc 600
acctacatca gcgtgggcac ctctacgctg aaccagcgtc tggtaccgcg tatcgcaacc 660
cgcagcaagg taaacggtca aagcggccgc atggaatttt tctggaccat cctgaaaccg 720
aacgacgcaa tcaacttcga atctaacggc aatttcatcg ctccggagta tgcgtacaaa 780
atcgtaaaga aaggtgatag cactatcatg aaatccgagc tggaatatgg caactgtaac 840
accaaatgcc agaccccgat gggtgcaatc aactcctcca tgccgtttca caacattcac 900

ccgctgacta tcggcgaatg tccgaaatac gttaaatcca atcgtctggt tctggctacc 960
ggtctgcgta actccccaca gcgtgaacgt cgtcgtaaga aacgtggtct gtttggcgcg 1020
atcgctggtt tcatcgaggg cggctggcag ggtatggttg atggctggta cggttatcat 1080
cattccaatg aacagggttc cggctacgcc gcagataaag aaagcactca gaaagcaatt 1140
gatggcgtaa ctaacaaagt aaattctatc attgataaaa tgaacaccca gttcgaggcg 1200
gttggtcgtg agttcaacaa cctggaacgc cgtatcgaaa acctgaacaa gaaaatggaa 1260
gacggttttc tggatgtgtg gacttacaat gctgaactgc tggtgctgat ggaaaacgag 1320
cgtaccctgg acttccacga cagcaacgtc aaaaatctgt atgacaaagt ccgtctgcag 1380
ctgcgtgata acgctaaaga gctgggtaat ggctgcttcg agttctatca caaatgcgac 1440
aacgaatgca tggaatctgt ccgcaacggc acttacgatt atccgcagta ctccgaagaa 1500
gcgcgcctga aacgcgaaga gatctccggt gtgaagctgg agtctattgg catctaccag 1560
atcctgtcca tctacagcac ctagtaagct gagcgcctac gcagc 1605

<210> 68
<211> 2193
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2.HA1-2 VN

<400> 68
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcagaa agcgtatgat 540
gtgaaagata cagcagtaac aacgaaagct tatgccaata atggtactac actggatgta 600
tcgggtcttg atgatgcagc tattaaagcg gctacgggtg gtacgaatgg tacggcttct 660
gtaaccggtg gtgcggttaa atttgacgca gataataaca agtactttgt tactattggt 720
ggctttactg gtgctgatgc cgccaaaaat ggcgattatg aagttaacgt tgctactgac 780
ggtacagtaa cccttgcggc tggcgcaact aaaaccacaa tgcctgctgg tgcgacaact 840
aaaacagaag tacaggagtt aaaagataca ccggcagttg tttcagcaga tgctaaaaat 900
gccttaattg ctggcggcgt tgacgctacc gatgctaatg gcgctgagtt ggtcaaaatg 960
tcttataccg ataaaaatgg taagacaatt gaaggcggtt atgcgcttaa agctggcgat 1020
aagtattacg ccgcagatta cgatgaagcg acaggagcaa ttaaagctaa aaccacaagt 1080
tatactgctg ctgacggcac taccaaaaca gcggctaacc aactgggtgg cgtagacggt 1140
aaaaccgaag tcgttactat cgacggtaaa acctacaatg ccagcaaagc cgctggtcat 1200
gatttcaaag cacaaccaga gctggcggaa gcagccgcta aaaccaccga aaacccgctg 1260
cagaaaattg atgccgcgct ggcgcaggtg gatgcgctgc gctctgatct gggtgcggta 1320
caaaaccgtt tcaactctgc tatcaccaac ctgggcaata ccgtaaacaa tctgtctgaa 1380
gcgcgtagcc gtatcgaaga ttccgactac gcgaccgaag tttccaacat gtctcgcgcg 1440
cagattttgc agcaggccgg tacttccgtt ctggcgcagg ctaaccaggt cccgcagaac 1500
gtgctgagcc tgttagcagg tgtgaaaccg ctgattctgc gcgattgctc tgtggcaggc 1560
tggctgctgg gcaaccctat gtgtgacgaa ttcattaacg ttccggaatg gtcttacatt 1620
gttgaaaaag ctaaccctgt caacgatctg tgttaccctg gtgactttaa cgattacgaa 1680
gaactgaagc acctgctgtc tcgtatcaat cacttcgaga aaatccagat catcccgaaa 1740
tcctcctgga gctcccacga agcttctctg ggcgtatcct ccgcgtgccc gtaccagggc 1800
aaatcctctt tctttcgtaa cgttgtttgg ctgatcaaga aaaactccac ctacccgacg 1860
atcaagcgta gctataataa caccaaccag gaagacctgc tggttctgtg gggcatccac 1920
catccaaacg atgctgcgga acagaccaag ctgtaccaga acccgaccac ctacatcagc 1980
gtgggcacct ctacgctgaa ccagcgtctg gtaccgcgta tcgcaacccg cagcaaggta 2040
aacggtcaaa gcggccgcat ggaatttttc tggaccatcc tgaaaccgaa cgacgcaatc 2100
aacttcgaat ctaacggcaa tttcatcgct ccggagtatg cgtacaaaat cgtaaagaaa 2160
ggtgatagca ctatcatgaa atccgagtag taa 2193

<210> 69
<211> 49
<212> DNA
<213> Artificial Sequence

<220>
<223> FOR PRIMER

<400> 69
agtcgctgag cctgttagca ggtgtgaaac cgctgattct gcgcgattg 49

<210> 70
<211> 48
<212> DNA
<213> Artificial Sequence

<220>
<223> REV PRIMER

<400> 70
tgacgctcag cttactactc ggatttcatg atagtgctat caccttc 48

<210> 71
<211> 2193
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2.HA1-2mut VN

<400> 71
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcagaa agcgtatgat 540
gtgaaagata cagcagtaac aacgaaagct tatgccaata atggtactac actggatgta 600
tcgggtcttg atgatgcagc tattaaagcg gctacgggtg gtacgaatgg tacggcttct 660
gtaaccggtg gtgcggttaa atttgacgca gataataaca agtactttgt tactattggt 720
ggctttactg gtgctgatgc cgccaaaaat ggcgattatg aagttaacgt tgctactgac 780
ggtacagtaa cccttgcggc tggcgcaact aaaaccacaa tgcctgctgg tgcgacaact 840
aaaacagaag tacaggagtt aaaagataca ccggcagttg tttcagcaga tgctaaaaat 900
gccttaattg ctggcggcgt tgacgctacc gatgctaatg gcgctgagtt ggtcaaaatg 960
tcttataccg ataaaaatgg taagacaatt gaaggcggtt atgcgcttaa agctggcgat 1020
aagtattacg ccgcagatta cgatgaagcg acaggagcaa ttaaagctaa aaccacaagt 1080
tatactgctg ctgacggcac taccaaaaca gcggctaacc aactgggtgg cgtagacggt 1140
aaaaccgaag tcgttactat cgacggtaaa acctacaatg ccagcaaagc cgctggtcat 1200
gatttcaaag cacaaccaga gctggcggaa gcagccgcta aaaccaccga aaacccgctg 1260
cagaaaattg atgccgcgct ggcgcaggtg gatgcgctgc gctctgatct gggtgcggta 1320
caaaaccgtt tcaactctgc tatcaccaac ctgggcaata ccgtaaacaa tctgtctgaa 1380
gcgcgtagcc gtatcgaaga ttccgactac gcgaccgaag tttccaacat gtctcgcgcg 1440
cagattttgc agcaggccgg tacttccgtt ctggcgcagg ctaaccaggt cccgcagaac 1500
gtgctgagcc tgttagcggg tgcgaaaccg ctgtctctgc gcgattgctc tgtggcaggc 1560
tggctgctgg gcaaccctat gtgtgacgaa ttcattaacg ttccggaatg gtctgatatt 1620

gccgaaaaag ctaaccctgt caacgatctg tgttaccctg gtgactttaa cgattacgaa 1680
gaactgaagc acctgctgtc tcgtatcaat cacttcgaga aaatccagat catcccgaaa 1740
tcctcctgga gctcccacga agcttctctg ggcgtatcct ccgcgtgccc gtaccagggc 1800
aaatcctctt tctttcgtaa cgttgtttgg ctgatcaaga aaaactccac ctacccgacg 1860
atcaagcgta gctataataa caccaaccag gaagacctgc tggttctgtg gggcatccac 1920
catccaaacg atgctgcgga acagaccaag ctgtaccaga acccgaccac ctacatcagc 1980
gtgggcacct ctacgctgaa ccagcgtctg gtaccgcgta tcgcaacccg cagcaaggta 2040
aacggtcaaa gcggccgcat ggaatttttc tggaccatcc tgaaaccgaa cgacgcaatc 2100
aacttcgaat ctaacggcaa tttcatcgct ccggagtatg cgtacaaaat cgtaaagaaa 2160
ggtgatagca ctatcatgaa atccgagtag taa 2193

<210> 72
<211> 675
<212> DNA
<213> Artificial Sequence

<220>
<223> 72 HA1-2mut VN

<400> 72
ggtgcgaaac cgctgtctct gcgcgattgc tctgtggcag gctggctgct gggcaaccct 60
atgtgtgacg aattcattaa cgttccggaa tggtctgata ttgccgaaaa agctaaccct 120
gtcaacgatc tgtgttaccc tggtgacttt aacgattacg aagaactgaa gcacctgctg 180
tctcgtatca atcacttcga gaaaatccag atcatcccga aatcctcctg gagctcccac 240
gaagcttctc tgggcgtatc ctccgcgtgc ccgtaccagg gcaaatcctc tttctttcgt 300
aacgttgttt ggctgatcaa gaaaaactcc acctacccga cgatcaagcg tagctataat 360
aacaccaacc aggaagacct gctggttctg tggggcatcc accatccaaa cgatgctgcg 420
gaacagacca agctgtacca gaacccgacc acctacatca gcgtgggcac ctctacgctg 480
aaccagcgtc tggtaccgcg tatcgcaacc cgcagcaagg taaacggtca aagcggccgc 540
atggaatttt tctggaccat cctgaaaccg aacgacgcaa tcaacttcga atctaacggc 600
aatttcatcg ctccggagta tgcgtacaaa atcgtaaaga aaggtgatag cactatcatg 660
aaatccgagt agtaa 675

<210> 73
<211> 2346
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2.HA1-1 IND

<400> 73
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcagaa agcgtatgat 540
gtgaaagata cagcagtaac aacgaaagct tatgccaata atggtactac actggatgta 600
tcgggtcttg atgatgcagc tattaaagcg gctacgggtg gtacgaatgg tacggcttct 660
gtaaccggtg gtgcggttaa atttgacgca gataataaca agtactttgt tactattggt 720
ggctttactg gtgctgatgc cgccaaaaat ggcgattatg aagttaacgt tgctactgac 780
ggtacagtaa cccttgcggc tggcgcaact aaaaccacaa tgcctgctgg tgcgacaact 840
aaaacagaag tacaggagtt aaaagataca ccggcagttg tttcagcaga tgctaaaaat 900
gccttaattg ctggcggcgt tgacgctacc gatgctaatg gcgctgagtt ggtcaaaatg 960
tcttataccg ataaaaatgg taagacaatt gaaggcggtt atgcgcttaa agctggcgat 1020

aagtattacg ccgcagatta cgatgaagcg acaggagcaa ttaaagctaa aaccacaagt 1080
tatactgctg ctgacggcac taccaaaaca gcggctaacc aactgggtgg cgtagacggt 1140
aaaaccgaag tcgttactat cgacggtaaa acctacaatg ccagcaaagc cgctggtcat 1200
gatttcaaag cacaaccaga gctggcggaa gcagccgcta aaaccaccga aaacccgctg 1260
cagaaaattg atgccgcgct ggcgcaggtg gatgcgctgc gctctgatct gggtgcggta 1320
caaaaccgtt tcaactctgc tatcaccaac ctgggcaata ccgtaaacaa tctgtctgaa 1380
gcgcgtagcc gtatcgaaga ttccgactac gcgaccgaag tttccaacat gtctcgcgcg 1440
cagattttgc agcaggccgg tacttccgtt ctggcgcagg ctaaccaggt cccgcagaac 1500
gtgctgtctc tgttagcgga gaaaacgcat aacggtaaac tgtgcgacct ggacggcgtg 1560
aagccgctga tcctgcgtga ctgttctgtt gctggctggc tgctgggcaa cccgatgtgc 1620
gatgagttca tcaacgttcc ggagtggtct tatattgttg aaaaagcgaa cccgaccaat 1680
gacctgtgct acccgggctc ttttaatgac tacgaggagc tgaaacatct gctgtctcgc 1740
atcaaccatt tcgaaaaaat tcagatcatc ccgaaatctt cctggagcga tcacgaagct 1800
tcttccggcg tatcttctgc atgcccttac ctgggttccc cgtctttctt ccgtaacgtg 1860
gtttggctga tcaagaagaa ctctacgtac ccgaccatca agaaatccta taacaacact 1920
aaccaggaag atctgctggt gctgtggggt atccatcatc caaacgatgc agccgaacag 1980
acccgcctgt accagaatcc gaccacgtac atctctatcg gcacttctac cctgaatcaa 2040
cgcctggtgc cgaaaatcgc aactcgcagc aaagtcaacg gccaatctgg tcgcatggag 2100
tttttctgga ccatcctgaa accgaacgac gcgattaact tcgaaagcaa cggcaacttc 2160
attgcaccgg aatatgctta caagatcgtt aaaaagggtg attctgcgat catgaaaagc 2220
gagctggaat acggcaactg caacactaaa tgccagaccc cgatgggtgc aatcaattcc 2280
agcatgccat ttcacaacat ccacccgctg accatcggcg agtgtccaaa atacgttaaa 2340
taatag 2346

<210> 74
<211> 46
<212> DNA
<213> Artificial Sequence

<220>
<223> FOR PRIMER

<400> 74
agtggctgag cctgttagcg gagaaaacgc ataacggtaa actgtg 46

<210> 75
<211> 1685
<212> DNA
<213> Artificial Sequence

<220>
<223> HA0s IND

<400> 75
gctgagcctg ttacgtgatc aaatctgcat cggttaccac gcaaacaact ccactgaaca 60
agtggatacg atcatggaaa agaacgtgac cgtgacccac gctcaagata tcctggagaa 120
aacgcataac ggtaaactgt gcgacctgga cggcgtgaag ccgctgatcc tgcgtgactg 180
ttctgttgct ggctggctgc tgggcaaccc gatgtgcgat gagttcatca acgttccgga 240
gtggtcttat attgttgaaa aagcgaaccc gaccaatgac ctgtgctacc cgggctcttt 300
taatgactac gaggagctga aacatctgct gtctcgcatc aaccatttcg aaaaaattca 360
gatcatcccg aaatcttcct ggagcgatca cgaagcttct tccggcgtat cttctgcatg 420
cccttacctg ggttccccgt ctttcttccg taacgtggtt tggctgatca agaagaactc 480
tacgtacccg accatcaaga aatcctataa caacactaac caggaagatc tgctggtgct 540
gtggggtatc catcatccaa acgatgcagc cgaacagacc cgcctgtacc agaatccgac 600
cacgtacatc tctatcggca cttctaccct gaatcaacgc ctggtgccga aaatcgcaac 660
tcgcagcaaa gtcaacggcc aatctggtcg catggagttt ttctggacca tcctgaaacc 720
gaacgacgcg attaacttcg aaagcaacgg caacttcatt gcaccggaat atgcttacaa 780
gatcgttaaa aagggtgatt ctgcgatcat gaaaagcgag ctggaatacg gcaactgcaa 840
cactaaatgc cagaccccga tgggtgcaat caattccagc atgccatttc acaacatcca 900

cccgctgacc atcggcgagt gtccaaaata cgttaaatcc aatcgtctgg tgctggcaac 960
gggcctgcgc aactctccgc aacgtgagtc tcgtcgcaaa aagcgtggcc tgtttggcgc 1020
aattgcgggc ttcatcgagg gcggctggca gggtatggtt gatggttggt acggttatca 1080
ccattccaac gagcagggct ctggttacgc tgcggacaaa gagtccacgc agaaagcgat 1140
cgatggcgtc accaacaaag tgaactctat catcgacaag atgaacactc agttcgaggc 1200
tgtgggccgt gaattcaaca atctggaacg tcgcatcgaa aacctgaaca aaaagatgga 1260
agacggtttc ctggacgtat ggacctataa cgcagagctg ctggttctga tggaaaacga 1320
acgtactctg gacttccacg attctaacgt taaaaacctg tacgacaaag ttcgtctgca 1380
gctgcgcgat aatgcaaaag aactgggtaa cggctgcttt gaattttacc acaaatgtga 1440
caacgaatgc atggaaagca tccgtaacgg tacctacaat tacccacagt actccgaaga 1500
agcgcgtctg aaacgtgaag aaatctccgg tgtaaaactg gaatccattg gcacgtatca 1560
gattctgtct atctactcca cggtagcgtc ctctctggcg ctggcaatta tgatggccgg 1620
cctgtctctg tggatgtgct ctaacggctc tctgcagtgc cgcatctgca tctaataggc 1680
tgagc 1685

<210> 76
<211> 2193
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2.HA1-2 IND

<400> 76
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcagaa agcgtatgat 540
gtgaaagata cagcagtaac aacgaaagct tatgccaata atggtactac actggatgta 600
tcgggtcttg atgatgcagc tattaaagcg gctacgggtg gtacgaatgg tacggcttct 660
gtaaccggtg gtgcggttaa atttgacgca gataataaca agtactttgt tactattggt 720
ggctttactg gtgctgatgc cgccaaaaat ggcgattatg aagttaacgt tgctactgac 780
ggtacagtaa cccttgcggc tggcgcaact aaaaccacaa tgcctgctgg tgcgacaact 840
aaaacagaag tacaggagtt aaaagataca ccggcagttg tttcagcaga tgctaaaaat 900
gccttaattg ctggcggcgt tgacgctacc gatgctaatg gcgctgagtt ggtcaaaatg 960
tcttataccg ataaaaatgg taagacaatt gaaggcggtt atgcgcttaa agctggcgat 1020
aagtattacg ccgcagatta cgatgaagcg acaggagcaa ttaaagctaa aaccacaagt 1080
tatactgctg ctgacggcac taccaaaaca gcggctaacc aactgggtgg cgtagacggt 1140
aaaaccgaag tcgttactat cgacggtaaa acctacaatg ccagcaaagc cgctggtcat 1200
gatttcaaag cacaaccaga gctggcggaa gcagccgcta aaaccaccga aaacccgctg 1260
cagaaaattg atgccgcgct ggcgcaggtg gatgcgctgc gctctgatct gggtgcggta 1320
caaaaccgtt tcaactctgc tatcaccaac ctgggcaata ccgtaaacaa tctgtctgaa 1380
gcgcgtagcc gtatcgaaga ttccgactac gcgaccgaag tttccaacat gtctcgcgcg 1440
cagattttgc agcaggccgg tacttccgtt ctggcgcagg ctaaccaggt cccgcagaac 1500
gtgctgagcc tgttagcagg tgtgaaaccg ctgattctgc gcgattgttc tgttgctggc 1560
tggctgctgg gcaacccgat gtgcgatgag ttcatcaacg ttccggagtg gtcttatatt 1620
gttgaaaaag cgaacccgac caatgacctg tgctacccgg gctcttttaa tgactacgag 1680
gagctgaaac atctgctgtc tcgcatcaac catttcgaaa aaattcagat catcccgaaa 1740
tcttcctgga gcgatcacga agcttcttcc ggcgtatctt ctgcatgccc ttacctgggt 1800
tccccgtctt tcttccgtaa cgtggtttgg ctgatcaaga agaactctac gtacccgacc 1860
atcaagaaat cctataacaa cactaaccag gaagatctgc tggtgctgtg gggtatccat 1920
catccaaacg atgcagccga acagacccgc ctgtaccaga atccgaccac gtacatctct 1980
atcggcactt ctaccctgaa tcaacgcctg gtgccgaaaa tcgcaactcg cagcaaagtc 2040
aacggccaat ctggtcgcat ggagtttttc tggaccatcc tgaaaccgaa cgacgcgatt 2100

aacttcgaaa gcaacggcaa cttcattgca ccggaatatg cttacaagat cgttaaaaag 2160
ggtgattctg cgatcatgaa aagcgagtaa tag 2193

<210> 77
<211> 43
<212> DNA
<213> Artificial Sequence

<220>
<223> REV PRIMER

<400> 77
agtcgctcag cctattactc gcttttcatg atcgcagaat cac 43

<210> 78
<211> 2193
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2.HA1-2mut IND

<400> 78
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcagaa agcgtatgat 540
gtgaaagata cagcagtaac aacgaaagct tatgccaata atggtactac actggatgta 600
tcgggtcttg atgatgcagc tattaaagcg gctacgggtg gtacgaatgg tacggcttct 660
gtaaccggtg gtgcggttaa atttgacgca gataataaca agtactttgt tactattggt 720
ggctttactg gtgctgatgc cgccaaaaat ggcgattatg aagttaacgt tgctactgac 780
ggtacagtaa cccttgcggc tggcgcaact aaaaccacaa tgcctgctgg tgcgacaact 840
aaaacagaag tacaggagtt aaaagataca ccggcagttg tttcagcaga tgctaaaaat 900
gccttaattg ctggcggcgt tgacgctacc gatgctaatg gcgctgagtt ggtcaaaatg 960
tcttataccg ataaaaatgg taagacaatt gaaggcggtt atgcgcttaa agctggcgat 1020
aagtattacg ccgcagatta cgatgaagcg acaggagcaa ttaaagctaa aaccacaagt 1080
tatactgctg ctgacggcac taccaaaaca gcggctaacc aactgggtgg cgtagacggt 1140
aaaaccgaag tcgttactat cgacggtaaa acctacaatg ccagcaaagc cgctggtcat 1200
gatttcaaag cacaaccaga gctggcggaa gcagccgcta aaaccaccga aaacccgctg 1260
cagaaaattg atgccgcgct ggcgcaggtg gatgcgctgc gctctgatct gggtgcggta 1320
caaaaccgtt tcaactctgc tatcaccaac ctgggcaata ccgtaaacaa tctgtctgaa 1380
gcgcgtagcc gtatcgaaga ttccgactac gcgaccgaag tttccaacat gtctcgcgcg 1440
cagattttgc agcaggccgg tacttccgtt ctggcgcagg ctaaccaggt cccgcagaac 1500
gtgctgtctc tgttagcggg cgcaaagccg ctgtctctgc gtgactgttc tgttgctggc 1560
tggctgctgg gcaacccgat gtgcgatgag ttcatcaacg ttccggagtg gtctgatatt 1620
gcggaaaaag cgaacccgac caatgacctg tgctacccgg gctcttttaa tgactacgag 1680
gagctgaaac atctgctgtc tcgcatcaac catttcgaaa aaattcagat catcccgaaa 1740
tcttcctgga gcgatcacga agcttcttcc ggcgtatctt ctgcatgccc ttacctgggt 1800
tccccgtctt tcttccgtaa cgtggtttgg ctgatcaaga agaactctac gtacccgacc 1860
atcaagaaat cctataacaa cactaaccag gaagatctgc tggtgctgtg gggtatccat 1920
catccaaacg atgcagccga acagacccgc ctgtaccaga atccgaccac gtacatctct 1980
atcggcactt ctaccctgaa tcaacgcctg gtgccgaaaa tcgcaactcg cagcaaagtc 2040
aacggccaat ctggtcgcat ggagtttttc tggaccatcc tgaaaccgaa cgacgcgatt 2100
aacttcgaaa gcaacggcaa cttcattgca ccggaatatg cttacaagat cgttaaaaag 2160

ggtgattctg cgatcatgaa aagcgagtaa tag 2193

<210> 79
<211> 682
<212> DNA
<213> Artificial Sequence

<220>
<223> HA1-2mut IND

<400> 79
ggcgcaaagc cgctgtctct gcgtgactgt tctgttgctg gctggctgct gggcaacccg 60
atgtgcgatg agttcatcaa cgttccggag tggtctgata ttgcggaaaa agcgaacccg 120
accaatgacc tgtgctaccc gggctctttt aatgactacg aggagctgaa acatctgctg 180
tctcgcatca accatttcga aaaaattcag atcatcccga aatcttcctg gagcgatcac 240
gaagcttctt ccggcgtatc ttctgcatgc ccttacctgg gttccccgtc tttcttccgt 300
aacgtggttt ggctgatcaa gaagaactct acgtacccga ccatcaagaa atcctataac 360
aacactaacc aggaagatct gctggtgctg tggggtatcc atcatccaaa cgatgcagcc 420
gaacagaccc gcctgtacca gaatccgacc acgtacatct ctatcggcac ttctaccctg 480
aatcaacgcc tggtgccgaa aatcgcaact cgcagcaaag tcaacggcca atctggtcgc 540
atggagtttt tctggaccat cctgaaaccg aacgacgcga ttaacttcga aagcaacggc 600
aacttcattg caccggaata tgcttacaag atcgttaaaa agggtgattc tgcgatcatg 660
aaaagcgagt aataggctga gc 682

<210> 80
<211> 2340
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2.HA1-1 NC

<400> 80
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcagaa agcgtatgat 540
gtgaaagata cagcagtaac aacgaaagct tatgccaata atggtactac actggatgta 600
tcgggtcttg atgatgcagc tattaaagcg gctacgggtg gtacgaatgg tacggcttct 660
gtaaccggtg gtgcggttaa atttgacgca gataataaca agtactttgt tactattggt 720
ggctttactg gtgctgatgc cgccaaaaat ggcgattatg aagttaacgt tgctactgac 780
ggtacagtaa cccttgcggc tggcgcaact aaaaccacaa tgcctgctgg tgcgacaact 840
aaaacagaag tacaggagtt aaaagataca ccggcagttg tttcagcaga tgctaaaaat 900
gccttaattg ctggcggcgt tgacgctacc gatgctaatg gcgctgagtt ggtcaaaatg 960
tcttataccg ataaaaatgg taagacaatt gaaggcggtt atgcgcttaa agctggcgat 1020
aagtattacg ccgcagatta cgatgaagcg acaggagcaa ttaaagctaa aaccacaagt 1080
tatactgctg ctgacggcac taccaaaaca gcggctaacc aactgggtgg cgtagacggt 1140
aaaaccgaag tcgttactat cgacggtaaa acctacaatg ccagcaaagc cgctggtcat 1200
gatttcaaag cacaaccaga gctggcggaa gcagccgcta aaaccaccga aaacccgctg 1260
cagaaaattg atgccgcgct ggcgcaggtg gatgcgctgc gctctgatct gggtgcggta 1320
caaaaccgtt tcaactctgc tatcaccaac ctgggcaata ccgtaaacaa tctgtctgaa 1380
gcgcgtagcc gtatcgaaga ttccgactac gcgaccgaag tttccaacat gtctcgcgcg 1440
cagattttgc agcaggccgg tacttccgtt ctggcgcagg ctaaccaggt cccgcagaac 1500
gtgctgagcc tgttagcgtc ccacaacggt aaactgtgtc tgctgaaagg catcgcaccg 1560

ctgcagctgg gtaactgtag cgttgcaggt tggatcctgg gtaacccgga gtgcgaactg 1620
ctgatttcca aggagagctg gtcttacatt gtcgaaacgc cgaatccgga aaacggcact 1680
tgttatccgg gttacttcgc cgattacgaa gaactgcgtg aacaactgtc ttccgtgagc 1740
tctttcgaac gtttcgagat cttcccgaaa gagagcagct ggccaaacca cactgttacc 1800
ggtgtgtctg cgagctgttc tcacaacggc aagtcttctt tctaccgcaa cctgctgtgg 1860
ctgacgggta agaatggcct gtatccgaac ctgtctaaat cttacgtaaa caacaaagag 1920
aaagaggtgc tggtcctgtg gggcgtacac catccaccaa atatcggcaa ccagcgcgcc 1980
ctgtaccaca ccgaaaacgc ttatgtgtcc gtggtgagct cccattacag ccgtcgtttt 2040
actccggaga ttgccaaacg tccgaaagtt cgtgatcagg aaggccgtat taactactac 2100
tggactctgc tggagccggg cgataccatc attttcgagg caaacggcaa cctgattgcg 2160
ccatggtacg cgttcgccct gagccgtggt tttggctccg gtattatcac ctctaacgcg 2220
ccaatggacg aatgcgacgc gaaatgccaa acgccgcagg gcgcaatcaa cagcagcctg 2280
ccgttccaga acgttcaccc ggttaccatc ggcgaatgcc ctaaatacgt gcgctaatag 2340

<210> 81
<211> 84
<212> DNA
<213> Artificial Sequence

<220>
<223> FOR PRIMER

<400> 81
aaccaggtcc cgcagaacgt gctgagcctg ttagcgtccc acaacggtaa actgtgtctg 60
ctgaaaggca tcgcaccgct gcag 84

<210> 82
<211> 78
<212> DNA
<213> Artificial Sequence

<220>
<223> REV PRIMER

<400> 82
gaacagctca gttcctacga gctcagccta ttagcgcacg tatttagggc attcgccgat 60
ggtaaccggg tgaacgtt 78

<210> 83
<211> 1653
<212> DNA
<213> Artificial Sequence

<220>
<223> HA0s NC

<400> 83
gacacgatct gtattggtta tcatgcaaac aactctactg acactgtaga tactgtgctg 60
gaaaagaacg taaccgttac ccacagcgtt aacctgctgg aagattccca caacggtaaa 120
ctgtgtctgc tgaaaggcat cgcaccgctg cagctgggta actgtagcgt tgcaggttgg 180
atcctgggta acccggagtg cgaactgctg atttccaagg agagctggtc ttacattgtc 240
gaaacgccga atccggaaaa cggcacttgt tatccgggtt acttcgccga ttacgaagaa 300
ctgcgtgaac aactgtcttc cgtgagctct ttcgaacgtt tcgagatctt cccgaaagag 360
agcagctggc caaaccacac tgttaccggt gtgtctgcga gctgttctca caacggcaag 420
tcttctttct accgcaacct gctgtggctg acgggtaaga atggcctgta tccgaacctg 480
tctaaatctt acgtaaacaa caaagagaaa gaggtgctgg tcctgtgggg cgtacaccat 540
ccaccaaata tcggcaacca gcgcgccctg taccacaccg aaaacgctta tgtgtccgtg 600
gtgagctccc attacagccg tcgttttact ccggagattg ccaaacgtcc gaaagttcgt 660

gatcaggaag gccgtattaa ctactactgg actctgctgg agccgggcga taccatcatt 720
ttcgaggcaa acggcaacct gattgcgcca tggtacgcgt tcgccctgag ccgtggtttt 780
ggctccggta ttatcacctc taacgcgcca atggacgaat gcgacgcgaa atgccaaacg 840
ccgcagggcg caatcaacag cagcctgccg ttccagaacg ttcacccggt taccatcggc 900
gaatgcccta aatacgtgcg ctccgccaaa ctgcgcatgg ttactggtct gcgtaacatc 960
ccgagcattc agtctcgcgg tctgttcggt gcgatcgcgg gcttcattga aggcggttgg 1020
accggtatgg tggatggttg gtacggctac catcaccaga acgaacaggg tagcggttac 1080
gctgccgacc agaaatccac ccagaacgct attaacggta tcaccaacaa agttaacagc 1140
gtaattgaga agatgaacac gcagttcacc gccgtaggta aggaattcaa caagctggaa 1200
cgtcgcatgg aaaacctgaa caaaaaggtg gacgacggct ttctggacat ctggacctac 1260
aacgctgaac tgctggtgct gctggaaaac gaacgtaccc tggatttcca cgactctaat 1320
gttaaaaacc tgtacgaaaa ggtcaagtct caactgaaaa acaatgcgaa ggaaatcggc 1380
aacggctgtt tcgaattcta ccataaatgc aacaacgaat gcatggaatc cgttaaaaac 1440
ggtacctatg actaccctaa atactccgaa gaaagcaaac tgaaccgcga gaaaatcgat 1500
ggtgtaaaac tggaatctat gggtgtttac cagatcctgg cgatctactc cacggtagcc 1560
agcagcctgg ttctgctggt tagcctgggt gcaattagct tctggatgtg ctctaacggc 1620
agcctgcaat gccgcatctg taataggctg agc 1653

<210> 84
<211> 2193
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2HA1-2 NC

<400> 84
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcagaa agcgtatgat 540
gtgaaagata cagcagtaac aacgaaagct tatgccaata atggtactac actggatgta 600
tcgggtcttg atgatgcagc tattaaagcg gctacgggtg gtacgaatgg tacggcttct 660
gtaaccggtg gtgcggttaa atttgacgca gataataaca agtactttgt tactattggt 720
ggctttactg gtgctgatgc cgccaaaaat ggcgattatg aagttaacgt tgctactgac 780
ggtacagtaa cccttgcggc tggcgcaact aaaaccacaa tgcctgctgg tgcgacaact 840
aaaacagaag tacaggagtt aaaagataca ccggcagttg tttcagcaga tgctaaaaat 900
gccttaattg ctggcggcgt tgacgctacc gatgctaatg gcgctgagtt ggtcaaaatg 960
tcttataccg ataaaaatgg taagacaatt gaaggcggtt atgcgcttaa agctggcgat 1020
aagtattacg ccgcagatta cgatgaagcg acaggagcaa ttaaagctaa aaccacaagt 1080
tatactgctg ctgacggcac taccaaaaca gcggctaacc aactgggtgg cgtagacggt 1140
aaaaccgaag tcgttactat cgacggtaaa acctacaatg ccagcaaagc cgctggtcat 1200
gatttcaaag cacaaccaga gctggcggaa gcagccgcta aaaccaccga aaacccgctg 1260
cagaaaattg atgccgcgct ggcgcaggtg gatgcgctgc gctctgatct gggtgcggta 1320
caaaaccgtt tcaactctgc tatcaccaac ctgggcaata ccgtaaacaa tctgtctgaa 1380
gcgcgtagcc gtatcgaaga ttccgactac gcgaccgaag tttccaacat gtctcgcgcg 1440
cagattttgc agcaggccgg tacttccgtt ctggcgcagg ctaaccaggt cccgcagaac 1500
gtgctgagcc tgttagcgaa aggcatcgca ccgctgcagc tgggtaactg tagcgttgca 1560
ggttggatcc tgggtaaccc ggagtgcgaa ctgctgattt ccaaggagag ctggtcttac 1620
attgtcgaaa cgccgaatcc ggaaaacggc acttgttatc cgggttactt cgccgattac 1680
gaagaactgc gtgaacaact gtcttccgtg agctctttcg aacgtttcga gatcttcccg 1740
aaagagagca gctggccaaa ccacactgtt accggtgtgt ctgcgagctg ttctcacaac 1800
ggcaagtctt ctttctaccg caacctgctg tggctgacgg gtaagaatgg cctgtatccg 1860
aacctgtcta aatcttacgt aaacaacaaa gagaaagagg tgctggtcct gtggggcgta 1920

caccatccac caaatatcgg caaccagcgc gccctgtacc acaccgaaaa cgcttatgtg 1980
tccgtggtga gctcccatta cagccgtcgt tttactccgg agattgccaa acgtccgaaa 2040
gttcgtgatc aggaaggccg tattaactac tactggactc tgctggagcc gggcgatacc 2100
atcattttcg aggcaaacgg caacctgatt gcgccatggt acgcgttcgc cctgagccgt 2160
ggttttggct ccggtattat cacctcttaa taa 2193

<210> 85
<211> 84
<212> DNA
<213> Artificial Sequence

<220>
<223> FOR PRIMER

<400> 85
aaccaggtcc cgcagaacgt gctgagcctg ttagcgaaag gcatcgcacc gctgcagctg 60
ggtaactgta gcgttgcagg ttgg 84

<210> 86
<211> 84
<212> DNA
<213> Artificial Sequence

<220>
<223> REV PRIMER

<400> 86
gaacagctca gttcctacga gctcagccta ttaagaggtg ataataccgg agccaaaacc 60
acggctcagg gcgaacgcgt acca 84

<210> 87
<211> 2193
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2.HA1-2mut NC

<400> 87
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcagaa agcgtatgat 540
gtgaaagata cagcagtaac aacgaaagct tatgccaata atggtactac actggatgta 600
tcgggtcttg atgatgcagc tattaaagcg gctacgggtg gtacgaatgg tacggcttct 660
gtaaccggtg gtgcggttaa atttgacgca gataataaca agtactttgt tactattggt 720
ggctttactg gtgctgatgc cgccaaaaat ggcgattatg aagttaacgt tgctactgac 780
ggtacagtaa cccttgcggc tggcgcaact aaaaccacaa tgcctgctgg tgcgacaact 840
aaaacagaag tacaggagtt aaaagataca ccggcagttg tttcagcaga tgctaaaaat 900
gccttaattg ctggcggcgt tgacgctacc gatgctaatg gcgctgagtt ggtcaaaatg 960
tcttataccg ataaaaatgg taagacaatt gaaggcggtt atgcgcttaa agctggcgat 1020
aagtattacg ccgcagatta cgatgaagcg acaggagcaa ttaaagctaa aaccacaagt 1080
tatactgctg ctgacggcac taccaaaaca gcggctaacc aactgggtgg cgtagacggt 1140
aaaaccgaag tcgttactat cgacggtaaa acctacaatg ccagcaaagc cgctggtcat 1200

gatttcaaag cacaaccaga gctggcggaa gcagccgcta aaaccaccga aaacccgctg 1260
cagaaaattg atgccgcgct ggcgcaggtg gatgcgctgc gctctgatct gggtgcggta 1320
caaaaccgtt tcaactctgc tatcaccaac ctgggcaata ccgtaaacaa tctgtctgaa 1380
gcgcgtagcc gtatcgaaga ttccgactac gcgaccgaag tttccaacat gtctcgcgcg 1440
cagattttgc agcaggccgg tacttccgtt ctggcgcagg ctaaccaggt cccgcagaac 1500
gtgctgagcc tgttagcgaa aggcgcggca ccgctgcagc tgggtaactg tagcgttgca 1560
ggttggatcc tgggtaaccc ggagtgcgaa ctgctgattt ccaaggagag ctggtctgat 1620
attgcagaaa cgccgaatcc ggaaaacggc acttgttatc cgggttactt cgccgattac 1680
gaagaactgc gtgaacaact gtcttccgtg agctctttcg aacgtttcga gatcttcccg 1740
aaagagagca gctggccaaa ccacactgtt accggtgtgt ctgcgagctg ttctcacaac 1800
ggcaagtctt ctttctaccg caacctgctg tggctgacgg gtaagaatgg cctgtatccg 1860
aacctgtcta aatcttacgt aaacaacaaa gagaaagagg tgctggtcct gtggggcgta 1920
caccatccac caaatatcgg caaccagcgc gccctgtacc acaccgaaaa cgcttatgtg 1980
tccgtggtga gctcccatta cagccgtcgt tttactccgg agattgccaa acgtccgaaa 2040
gttcgtgatc aggaaggccg tattaactac tactggactc tgctggagcc gggcgatacc 2100
atcattttcg aggcaaacgg caacctgatt gcgccatggt acgcgttcgc cctgagccgt 2160
ggttttggct ccggtattat cacctcttaa tag 2193

<210> 88
<211> 698
<212> DNA
<213> Artificial Sequence

<220>
<223> HA1-2mut NC

<400> 88
gctgagcctg ttagcgaaag gcgcggcacc gctgcagctg ggtaactgta gcgttgcagg 60
ttggatcctg ggtaacccgg agtgcgaact gctgatttcc aaggagagct ggtctgatat 120
tgcagaaacg ccgaatccgg aaaacggcac ttgttatccg ggttacttcg ccgattacga 180
agaactgcgt gaacaactgt cttccgtgag ctctttcgaa cgtttcgaga tcttcccgaa 240
agagagcagc tggccaaacc acactgttac cggtgtgtct gcgagctgtt ctcacaacgg 300
caagtcttct ttctaccgca acctgctgtg gctgacgggt aagaatggcc tgtatccgaa 360
cctgtctaaa tcttacgtaa acaacaaaga gaaagaggtg ctggtcctgt ggggcgtaca 420
ccatccacca aatatcggca accagcgcgc cctgtaccac accgaaaacg cttatgtgtc 480
cgtggtgagc tcccattaca gccgtcgttt tactccggag attgccaaac gtccgaaagt 540
tcgtgatcag gaaggccgta ttaactacta ctggactctg ctggagccgg gcgataccat 600
cattttcgag gcaaacggca acctgattgc gccatggtac gcgttcgccc tgagccgtgg 660
ttttggctcc ggtattatca cctcttaata ggctgagc 698

<210> 89
<211> 784
<212> PRT
<213> Artificial Sequence

<220>
<223> STF2.HA1-1his (PR8)

<400> 89
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
1 5 10 15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
20 25 30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
35 40 45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
50 55 60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65 70 75 80

```
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
                85                  90                  95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
            100                 105                 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
        115                 120                 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
    130                 135                 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145                 150                 155                 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val Gln
                165                 170                 175
Lys Ala Tyr Asp Val Lys Asp Thr Ala Val Thr Thr Lys Ala Tyr Ala
            180                 185                 190
Asn Asn Gly Thr Thr Leu Asp Val Ser Gly Leu Asp Asp Ala Ala Ile
        195                 200                 205
Lys Ala Ala Thr Gly Gly Thr Asn Gly Thr Ala Ser Val Thr Gly Gly
    210                 215                 220
Ala Val Lys Phe Asp Ala Asp Asn Asn Lys Tyr Phe Val Thr Ile Gly
225                 230                 235                 240
Gly Phe Thr Gly Ala Asp Ala Ala Lys Asn Gly Asp Tyr Glu Val Asn
                245                 250                 255
Val Ala Thr Asp Gly Thr Val Thr Leu Ala Ala Gly Ala Thr Lys Thr
            260                 265                 270
Thr Met Pro Ala Gly Ala Thr Thr Lys Thr Glu Val Gln Glu Leu Lys
        275                 280                 285
Asp Thr Pro Ala Val Val Ser Ala Asp Ala Lys Asn Ala Leu Ile Ala
    290                 295                 300
Gly Gly Val Asp Ala Thr Asp Ala Asn Gly Ala Glu Leu Val Lys Met
305                 310                 315                 320
Ser Tyr Thr Asp Lys Asn Gly Lys Thr Ile Glu Gly Gly Tyr Ala Leu
                325                 330                 335
Lys Ala Gly Asp Lys Tyr Tyr Ala Ala Asp Tyr Asp Glu Ala Thr Gly
            340                 345                 350
Ala Ile Lys Ala Lys Thr Thr Ser Tyr Thr Ala Ala Asp Gly Thr Thr
        355                 360                 365
Lys Thr Ala Ala Asn Gln Leu Gly Gly Val Asp Gly Lys Thr Glu Val
    370                 375                 380
Val Thr Ile Asp Gly Lys Thr Tyr Asn Ala Ser Lys Ala Ala Gly His
385                 390                 395                 400
Asp Phe Lys Ala Gln Pro Glu Leu Ala Glu Ala Ala Ala Lys Thr Thr
                405                 410                 415
Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala
            420                 425                 430
Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile
        435                 440                 445
Thr Asn Leu Gly Asn Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg
    450                 455                 460
Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala
465                 470                 475                 480
Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln
                485                 490                 495
Val Pro Gln Asn Val Leu Ser Leu Leu Ala Ser His Asn Gly Lys Leu
            500                 505                 510
Cys Arg Leu Lys Gly Ile Ala Pro Leu Gln Leu Gly Lys Cys Asn Ile
        515                 520                 525
Ala Gly Trp Leu Leu Gly Asn Pro Glu Cys Asp Pro Leu Leu Pro Val
    530                 535                 540
Arg Ser Trp Ser Tyr Ile Val Glu Thr Pro Asn Ser Glu Asn Gly Ile
```

```
545                 550                 555                 560
Cys Tyr Pro Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu
                565                 570                 575
Ser Ser Val Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Glu Ser
            580                 585                 590
Ser Trp Pro Asn His Asn Thr Asn Gly Val Thr Ala Ala Cys Ser His
        595                 600                 605
Glu Gly Lys Ser Ser Phe Tyr Arg Asn Leu Leu Trp Leu Thr Glu Lys
    610                 615                 620
Glu Gly Ser Tyr Pro Lys Leu Lys Asn Ser Tyr Val Asn Lys Lys Gly
625                 630                 635                 640
Lys Glu Val Leu Val Leu Trp Gly Ile His His Pro Pro Asn Ser Lys
                645                 650                 655
Glu Gln Gln Asn Leu Tyr Gln Asn Glu Asn Ala Tyr Val Ser Val Val
            660                 665                 670
Thr Ser Asn Tyr Asn Arg Arg Phe Thr Pro Glu Ile Ala Glu Arg Pro
        675                 680                 685
Lys Val Arg Asp Gln Ala Gly Arg Met Asn Tyr Tyr Trp Thr Leu Leu
    690                 695                 700
Lys Pro Gly Asp Thr Ile Ile Phe Glu Ala Asn Gly Asn Leu Ile Ala
705                 710                 715                 720
Pro Met Tyr Ala Phe Ala Leu Ser Arg Gly Phe Gly Ser Gly Ile Ile
                725                 730                 735
Thr Ser Asn Ala Ser Met His Glu Cys Asn Thr Lys Cys Gln Thr Pro
            740                 745                 750
Leu Gly Ala Ile Asn Ser Ser Leu Pro Tyr Gln Asn Ile His Pro Val
        755                 760                 765
Thr Ile Gly Glu Cys Pro Lys Tyr Val Arg His His His His His His
    770                 775                 780


<210> 90
<211> 729
<212> PRT
<213> Artificial Sequence

<220>
<223> STF2.HA1-2 (PR8)

<400> 90
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
 1               5                  10                  15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
            20                  25                  30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
        35                  40                  45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
    50                  55                  60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65                  70                  75                  80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
                85                  90                  95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
            100                 105                 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
        115                 120                 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
    130                 135                 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
```

```
145                 150                 155                 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val Gln
                165                 170                 175
Lys Ala Tyr Asp Val Lys Asp Thr Ala Val Thr Thr Lys Ala Tyr Ala
            180                 185                 190
Asn Asn Gly Thr Thr Leu Asp Val Ser Gly Leu Asp Asp Ala Ala Ile
        195                 200                 205
Lys Ala Ala Thr Gly Gly Thr Asn Gly Thr Ala Ser Val Thr Gly Gly
    210                 215                 220
Ala Val Lys Phe Asp Ala Asp Asn Asn Lys Tyr Phe Val Thr Ile Gly
225                 230                 235                 240
Gly Phe Thr Gly Ala Asp Ala Ala Lys Asn Gly Asp Tyr Glu Val Asn
                245                 250                 255
Val Ala Thr Asp Gly Thr Val Thr Leu Ala Ala Gly Ala Thr Lys Thr
            260                 265                 270
Thr Met Pro Ala Gly Ala Thr Thr Lys Thr Glu Val Gln Glu Leu Lys
        275                 280                 285
Asp Thr Pro Ala Val Val Ser Ala Asp Ala Lys Asn Ala Leu Ile Ala
    290                 295                 300
Gly Gly Val Asp Ala Thr Asp Ala Asn Gly Ala Glu Leu Val Lys Met
305                 310                 315                 320
Ser Tyr Thr Asp Lys Asn Gly Lys Thr Ile Glu Gly Gly Tyr Ala Leu
                325                 330                 335
Lys Ala Gly Asp Lys Tyr Tyr Ala Ala Asp Tyr Asp Glu Ala Thr Gly
            340                 345                 350
Ala Ile Lys Ala Lys Thr Thr Ser Tyr Thr Ala Ala Asp Gly Thr Thr
        355                 360                 365
Lys Thr Ala Ala Asn Gln Leu Gly Gly Val Asp Gly Lys Thr Glu Val
    370                 375                 380
Val Thr Ile Asp Gly Lys Thr Tyr Asn Ala Ser Lys Ala Ala Gly His
385                 390                 395                 400
Asp Phe Lys Ala Gln Pro Glu Leu Ala Glu Ala Ala Ala Lys Thr Thr
                405                 410                 415
Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala
            420                 425                 430
Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile
        435                 440                 445
Thr Asn Leu Gly Asn Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg
    450                 455                 460
Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala
465                 470                 475                 480
Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln
                485                 490                 495
Val Pro Gln Asn Val Leu Ser Leu Leu Ala Lys Gly Ile Ala Pro Leu
            500                 505                 510
Gln Leu Gly Lys Cys Asn Ile Ala Gly Trp Leu Leu Gly Asn Pro Glu
        515                 520                 525
Cys Asp Pro Leu Leu Pro Val Arg Ser Trp Ser Tyr Ile Val Glu Thr
    530                 535                 540
Pro Asn Ser Glu Asn Gly Ile Cys Tyr Pro Gly Asp Phe Ile Asp Tyr
545                 550                 555                 560
Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe Glu Arg Phe
                565                 570                 575
Glu Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His Asn Thr Asn Gly
            580                 585                 590
Val Thr Ala Ala Cys Ser His Glu Gly Lys Ser Ser Phe Tyr Arg Asn
        595                 600                 605
Leu Leu Trp Leu Thr Glu Lys Glu Gly Ser Tyr Pro Lys Leu Lys Asn
    610                 615                 620
```

```
Ser Tyr Val Asn Lys Lys Gly Lys Glu Val Leu Val Leu Trp Gly Ile
625                 630                 635                 640
His His Pro Pro Asn Ser Lys Glu Gln Gln Asn Leu Tyr Gln Asn Glu
                645                 650                 655
Asn Ala Tyr Val Ser Val Val Thr Ser Asn Tyr Asn Arg Arg Phe Thr
            660                 665                 670
Pro Glu Ile Ala Glu Arg Pro Lys Val Arg Asp Gln Ala Gly Arg Met
        675                 680                 685
Asn Tyr Tyr Trp Thr Leu Leu Lys Pro Gly Asp Thr Ile Ile Phe Glu
    690                 695                 700
Ala Asn Gly Asn Leu Ile Ala Pro Met Tyr Ala Phe Ala Leu Ser Arg
705                 710                 715                 720
Gly Phe Gly Ser Gly Ile Ile Thr Ser
                725
```

<210> 91
<211> 729
<212> PRT
<213> Artificial Sequence

<220>
<223> STF2.HA1-2mut (PR8)

<400> 91

```
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
 1               5                  10                  15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
            20                  25                  30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
        35                  40                  45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
    50                  55                  60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65                  70                  75                  80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
                85                  90                  95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
            100                 105                 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
        115                 120                 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
    130                 135                 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145                 150                 155                 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val Gln
                165                 170                 175
Lys Ala Tyr Asp Val Lys Asp Thr Ala Val Thr Thr Lys Ala Tyr Ala
            180                 185                 190
Asn Asn Gly Thr Thr Leu Asp Val Ser Gly Leu Asp Asp Ala Ala Ile
        195                 200                 205
Lys Ala Ala Thr Gly Gly Thr Asn Gly Thr Ala Ser Val Thr Gly Gly
    210                 215                 220
Ala Val Lys Phe Asp Ala Asp Asn Asn Lys Tyr Phe Val Thr Ile Gly
225                 230                 235                 240
Gly Phe Thr Gly Ala Asp Ala Ala Lys Asn Gly Asp Tyr Glu Val Asn
                245                 250                 255
Val Ala Thr Asp Gly Thr Val Thr Leu Ala Ala Gly Ala Thr Lys Thr
            260                 265                 270
```

```
Thr Met Pro Ala Gly Ala Thr Thr Lys Thr Glu Val Gln Glu Leu Lys
        275                 280                 285
Asp Thr Pro Ala Val Val Ser Ala Asp Ala Lys Asn Ala Leu Ile Ala
    290                 295                 300
Gly Gly Val Asp Ala Thr Asp Ala Asn Gly Ala Glu Leu Val Lys Met
305                 310                 315                 320
Ser Tyr Thr Asp Lys Asn Gly Lys Thr Ile Glu Gly Gly Tyr Ala Leu
                325                 330                 335
Lys Ala Gly Asp Lys Tyr Tyr Ala Ala Asp Tyr Asp Glu Ala Thr Gly
            340                 345                 350
Ala Ile Lys Ala Lys Thr Thr Ser Tyr Thr Ala Ala Asp Gly Thr Thr
        355                 360                 365
Lys Thr Ala Ala Asn Gln Leu Gly Gly Val Asp Gly Lys Thr Glu Val
    370                 375                 380
Val Thr Ile Asp Gly Lys Thr Tyr Asn Ala Ser Lys Ala Ala Gly His
385                 390                 395                 400
Asp Phe Lys Ala Gln Pro Glu Leu Ala Glu Ala Ala Ala Lys Thr Thr
                405                 410                 415
Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala
            420                 425                 430
Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile
        435                 440                 445
Thr Asn Leu Gly Asn Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg
    450                 455                 460
Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala
465                 470                 475                 480
Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln
                485                 490                 495
Val Pro Gln Asn Val Leu Ser Leu Leu Ala Lys Gly Ala Ala Pro Leu
            500                 505                 510
Gln Leu Gly Lys Cys Asn Ile Ala Gly Trp Leu Leu Gly Asn Pro Glu
        515                 520                 525
Cys Asp Pro Leu Leu Pro Val Arg Ser Trp Ser Asp Ile Ala Glu Thr
    530                 535                 540
Pro Asn Ser Glu Asn Gly Ile Cys Tyr Pro Gly Asp Phe Ile Asp Tyr
545                 550                 555                 560
Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe Glu Arg Phe
                565                 570                 575
Glu Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His Asn Thr Asn Gly
            580                 585                 590
Val Thr Ala Ala Cys Ser His Glu Gly Lys Ser Ser Phe Tyr Arg Asn
        595                 600                 605
Leu Leu Trp Leu Thr Glu Lys Glu Gly Ser Tyr Pro Lys Leu Lys Asn
    610                 615                 620
Ser Tyr Val Asn Lys Lys Gly Lys Glu Val Leu Val Leu Trp Gly Ile
625                 630                 635                 640
His His Pro Pro Asn Ser Lys Glu Gln Gln Asn Leu Tyr Gln Asn Glu
                645                 650                 655
Asn Ala Tyr Val Ser Val Val Thr Ser Asn Tyr Asn Arg Arg Phe Thr
            660                 665                 670
Pro Glu Ile Ala Glu Arg Pro Lys Val Arg Asp Gln Ala Gly Arg Met
        675                 680                 685
Asn Tyr Tyr Trp Thr Leu Leu Lys Pro Gly Asp Thr Ile Ile Phe Glu
    690                 695                 700
Ala Asn Gly Asn Leu Ile Ala Pro Met Tyr Ala Phe Ala Leu Ser Arg
705                 710                 715                 720
Gly Phe Gly Ser Gly Ile Ile Thr Ser
                725
```

<210> 92
<211> 682
<212> PRT
<213> Artificial Sequence

<220>
<223> STF2.HA1-3 (PR8)

<400> 92
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
1 5 10 15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
20 25 30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
35 40 45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
50 55 60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65 70 75 80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
85 90 95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
100 105 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
115 120 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
130 135 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145 150 155 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val Gln
165 170 175
Lys Ala Tyr Asp Val Lys Asp Thr Ala Val Thr Thr Lys Ala Tyr Ala
180 185 190
Asn Asn Gly Thr Thr Leu Asp Val Ser Gly Leu Asp Asp Ala Ala Ile
195 200 205
Lys Ala Ala Thr Gly Gly Thr Asn Gly Thr Ala Ser Val Thr Gly Gly
210 215 220
Ala Val Lys Phe Asp Ala Asp Asn Asn Lys Tyr Phe Val Thr Ile Gly
225 230 235 240
Gly Phe Thr Gly Ala Asp Ala Ala Lys Asn Gly Asp Tyr Glu Val Asn
245 250 255
Val Ala Thr Asp Gly Thr Val Thr Leu Ala Ala Gly Ala Thr Lys Thr
260 265 270
Thr Met Pro Ala Gly Ala Thr Thr Lys Thr Glu Val Gln Glu Leu Lys
275 280 285
Asp Thr Pro Ala Val Val Ser Ala Asp Ala Lys Asn Ala Leu Ile Ala
290 295 300
Gly Gly Val Asp Ala Thr Asp Ala Asn Gly Ala Glu Leu Val Lys Met
305 310 315 320
Ser Tyr Thr Asp Lys Asn Gly Lys Thr Ile Glu Gly Gly Tyr Ala Leu
325 330 335
Lys Ala Gly Asp Lys Tyr Tyr Ala Ala Asp Tyr Asp Glu Ala Thr Gly
340 345 350
Ala Ile Lys Ala Lys Thr Thr Ser Tyr Thr Ala Ala Asp Gly Thr Thr
355 360 365
Lys Thr Ala Ala Asn Gln Leu Gly Gly Val Asp Gly Lys Thr Glu Val
370 375 380
Val Thr Ile Asp Gly Lys Thr Tyr Asn Ala Ser Lys Ala Ala Gly His

```
385                 390                 395                 400
Asp Phe Lys Ala Gln Pro Glu Leu Ala Glu Ala Ala Ala Lys Thr Thr
                405                 410                 415
Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala
            420                 425                 430
Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile
        435                 440                 445
Thr Asn Leu Gly Asn Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg
    450                 455                 460
Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala
465                 470                 475                 480
Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln
                485                 490                 495
Val Pro Gln Asn Val Leu Ser Leu Leu Ala Asn Ser Glu Asn Gly Ile
            500                 505                 510
Cys Tyr Pro Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu
        515                 520                 525
Ser Ser Val Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Glu Ser
    530                 535                 540
Ser Trp Pro Asn His Asn Thr Asn Gly Val Thr Ala Ala Cys Ser His
545                 550                 555                 560
Glu Gly Lys Ser Ser Phe Tyr Arg Asn Leu Leu Trp Leu Thr Glu Lys
                565                 570                 575
Glu Gly Ser Tyr Pro Lys Leu Lys Asn Ser Tyr Val Asn Lys Lys Gly
            580                 585                 590
Lys Glu Val Leu Val Leu Trp Gly Ile His His Pro Pro Asn Ser Lys
        595                 600                 605
Glu Gln Gln Asn Leu Tyr Gln Asn Glu Asn Ala Tyr Val Ser Val Val
    610                 615                 620
Thr Ser Asn Tyr Asn Arg Arg Phe Thr Pro Glu Ile Ala Glu Arg Pro
625                 630                 635                 640
Lys Val Arg Asp Gln Ala Gly Arg Met Asn Tyr Tyr Trp Thr Leu Leu
                645                 650                 655
Lys Pro Gly Asp Thr Ile Ile Phe Glu Ala Asn Gly Asn Leu Ile Ala
            660                 665                 670
Pro Met Tyr Ala Phe Ala Leu Ser Arg Gly
        675                 680


<210> 93
<211> 229
<212> PRT
<213> Artificial Sequence

<220>
<223> HA1-2his (PR8)

<400> 93
Lys Gly Ile Ala Pro Leu Gln Leu Gly Lys Cys Asn Ile Ala Gly Trp
 1               5                  10                  15
Leu Leu Gly Asn Pro Glu Cys Asp Pro Leu Leu Pro Val Arg Ser Trp
            20                  25                  30
Ser Tyr Ile Val Glu Thr Pro Asn Ser Glu Asn Gly Ile Cys Tyr Pro
        35                  40                  45
Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val
    50                  55                  60
Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Glu Ser Ser Trp Pro
65                  70                  75                  80
Asn His Asn Thr Asn Gly Val Thr Ala Ala Cys Ser His Glu Gly Lys
```

```
                85                  90                  95
Ser Ser Phe Tyr Arg Asn Leu Leu Trp Leu Thr Glu Lys Glu Gly Ser
            100                 105                 110
Tyr Pro Lys Leu Lys Asn Ser Tyr Val Asn Lys Lys Gly Lys Glu Val
        115                 120                 125
Leu Val Leu Trp Gly Ile His His Pro Pro Asn Ser Lys Glu Gln Gln
    130                 135                 140
Asn Leu Tyr Gln Asn Glu Asn Ala Tyr Val Ser Val Val Thr Ser Asn
145                 150                 155                 160
Tyr Asn Arg Arg Phe Thr Pro Glu Ile Ala Glu Arg Pro Lys Val Arg
                165                 170                 175
Asp Gln Ala Gly Arg Met Asn Tyr Tyr Trp Thr Leu Leu Lys Pro Gly
            180                 185                 190
Asp Thr Ile Ile Phe Glu Ala Asn Gly Asn Leu Ile Ala Pro Met Tyr
        195                 200                 205
Ala Phe Ala Leu Ser Arg Gly Phe Gly Ser Gly Ile Ile Thr Ser His
    210                 215                 220
His His His His His
225


<210> 94
<211> 611
<212> PRT
<213> Artificial Sequence

<220>
<223> STF2.4xM2e

<400> 94
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
 1                5                  10                  15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
            20                  25                  30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
        35                  40                  45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
    50                  55                  60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65                  70                  75                  80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
                85                  90                  95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
            100                 105                 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
        115                 120                 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
    130                 135                 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145                 150                 155                 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val Gln
                165                 170                 175
Lys Ala Tyr Asp Val Lys Asp Thr Ala Val Thr Thr Lys Ala Tyr Ala
            180                 185                 190
Asn Asn Gly Thr Thr Leu Asp Val Ser Gly Leu Asp Asp Ala Ala Ile
        195                 200                 205
Lys Ala Ala Thr Gly Gly Thr Asn Gly Thr Ala Ser Val Thr Gly Gly
    210                 215                 220
Ala Val Lys Phe Asp Ala Asp Asn Asn Lys Tyr Phe Val Thr Ile Gly
```

```
225                 230                 235                 240
Gly Phe Thr Gly Ala Asp Ala Ala Lys Asn Gly Asp Tyr Glu Val Asn
                245                 250                 255
Val Ala Thr Asp Gly Thr Val Thr Leu Ala Ala Gly Ala Thr Lys Thr
            260                 265                 270
Thr Met Pro Ala Gly Ala Thr Thr Lys Thr Glu Val Gln Glu Leu Lys
        275                 280                 285
Asp Thr Pro Ala Val Val Ser Ala Asp Ala Lys Asn Ala Leu Ile Ala
    290                 295                 300
Gly Gly Val Asp Ala Thr Asp Ala Asn Gly Ala Glu Leu Val Lys Met
305                 310                 315                 320
Ser Tyr Thr Asp Lys Asn Gly Lys Thr Ile Glu Gly Gly Tyr Ala Leu
                325                 330                 335
Lys Ala Gly Asp Lys Tyr Tyr Ala Ala Asp Tyr Asp Glu Ala Thr Gly
            340                 345                 350
Ala Ile Lys Ala Lys Thr Thr Ser Tyr Thr Ala Ala Asp Gly Thr Thr
        355                 360                 365
Lys Thr Ala Ala Asn Gln Leu Gly Gly Val Asp Gly Lys Thr Glu Val
    370                 375                 380
Val Thr Ile Asp Gly Lys Thr Tyr Asn Ala Ser Lys Ala Ala Gly His
385                 390                 395                 400
Asp Phe Lys Ala Gln Pro Glu Leu Ala Glu Ala Ala Ala Lys Thr Thr
                405                 410                 415
Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala
            420                 425                 430
Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile
        435                 440                 445
Thr Asn Leu Gly Asn Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg
    450                 455                 460
Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala
465                 470                 475                 480
Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln
                485                 490                 495
Val Pro Gln Asn Val Leu Ser Leu Leu Arg Leu Ser Leu Leu Thr Glu
            500                 505                 510
Val Glu Thr Pro Ile Arg Asn Glu Trp Gly Ser Arg Ser Asn Asp Ser
        515                 520                 525
Ser Asp Pro Leu Glu Ser Leu Leu Thr Glu Val Glu Thr Pro Ile Arg
    530                 535                 540
Asn Glu Trp Gly Ser Arg Ser Asn Asp Ser Ser Asp Pro Gly Ser Ser
545                 550                 555                 560
Leu Leu Thr Glu Val Glu Thr Pro Ile Arg Asn Glu Trp Gly Ser Arg
                565                 570                 575
Ser Asn Asp Ser Ser Asp Pro Glu Leu Ser Leu Leu Thr Glu Val Glu
            580                 585                 590
Thr Pro Ile Arg Asn Glu Trp Gly Ser Arg Ser Asn Asp Ser Ser Asp
        595                 600                 605
Pro Ser Arg
    610
```

<210> 95
<211> 729
<212> PRT
<213> Artificial Sequence

<220>
<223> STF2.HA1-2 (VN)

<400> 95
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
1 5 10 15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
20 25 30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
35 40 45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
50 55 60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65 70 75 80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
85 90 95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
100 105 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
115 120 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
130 135 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145 150 155 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val Gln
165 170 175
Lys Ala Tyr Asp Val Lys Asp Thr Ala Val Thr Thr Lys Ala Tyr Ala
180 185 190
Asn Asn Gly Thr Thr Leu Asp Val Ser Gly Leu Asp Asp Ala Ala Ile
195 200 205
Lys Ala Ala Thr Gly Gly Thr Asn Gly Thr Ala Ser Val Thr Gly Gly
210 215 220
Ala Val Lys Phe Asp Ala Asp Asn Asn Lys Tyr Phe Val Thr Ile Gly
225 230 235 240
Gly Phe Thr Gly Ala Asp Ala Ala Lys Asn Gly Asp Tyr Glu Val Asn
245 250 255
Val Ala Thr Asp Gly Thr Val Thr Leu Ala Ala Gly Ala Thr Lys Thr
260 265 270
Thr Met Pro Ala Gly Ala Thr Thr Lys Thr Glu Val Gln Glu Leu Lys
275 280 285
Asp Thr Pro Ala Val Val Ser Ala Asp Ala Lys Asn Ala Leu Ile Ala
290 295 300
Gly Gly Val Asp Ala Thr Asp Ala Asn Gly Ala Glu Leu Val Lys Met
305 310 315 320
Ser Tyr Thr Asp Lys Asn Gly Lys Thr Ile Glu Gly Gly Tyr Ala Leu
325 330 335
Lys Ala Gly Asp Lys Tyr Tyr Ala Ala Asp Tyr Asp Glu Ala Thr Gly
340 345 350
Ala Ile Lys Ala Lys Thr Thr Ser Tyr Thr Ala Ala Asp Gly Thr Thr
355 360 365
Lys Thr Ala Ala Asn Gln Leu Gly Gly Val Asp Gly Lys Thr Glu Val
370 375 380
Val Thr Ile Asp Gly Lys Thr Tyr Asn Ala Ser Lys Ala Ala Gly His
385 390 395 400
Asp Phe Lys Ala Gln Pro Glu Leu Ala Glu Ala Ala Ala Lys Thr Thr
405 410 415
Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala
420 425 430
Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile
435 440 445
Thr Asn Leu Gly Asn Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg
450 455 460

```
Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala
465                 470                 475                 480
Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln
                485                 490                 495
Val Pro Gln Asn Val Leu Ser Leu Leu Ala Gly Val Lys Pro Leu Ile
            500                 505                 510
Leu Arg Asp Cys Ser Val Ala Gly Trp Leu Leu Gly Asn Pro Met Cys
        515                 520                 525
Asp Glu Phe Ile Asn Val Pro Glu Trp Ser Tyr Ile Val Glu Lys Ala
    530                 535                 540
Asn Pro Val Asn Asp Leu Cys Tyr Pro Gly Asp Phe Asn Asp Tyr Glu
545                 550                 555                 560
Glu Leu Lys His Leu Leu Ser Arg Ile Asn His Phe Glu Lys Ile Gln
                565                 570                 575
Ile Ile Pro Lys Ser Ser Trp Ser Ser His Glu Ala Ser Leu Gly Val
            580                 585                 590
Ser Ser Ala Cys Pro Tyr Gln Gly Lys Ser Ser Phe Phe Arg Asn Val
        595                 600                 605
Val Trp Leu Ile Lys Lys Asn Ser Thr Tyr Pro Thr Ile Lys Arg Ser
    610                 615                 620
Tyr Asn Asn Thr Asn Gln Glu Asp Leu Leu Val Leu Trp Gly Ile His
625                 630                 635                 640
His Pro Asn Asp Ala Ala Glu Gln Thr Lys Leu Tyr Gln Asn Pro Thr
                645                 650                 655
Thr Tyr Ile Ser Val Gly Thr Ser Thr Leu Asn Gln Arg Leu Val Pro
            660                 665                 670
Arg Ile Ala Thr Arg Ser Lys Val Asn Gly Gln Ser Gly Arg Met Glu
        675                 680                 685
Phe Phe Trp Thr Ile Leu Lys Pro Asn Asp Ala Ile Asn Phe Glu Ser
    690                 695                 700
Asn Gly Asn Phe Ile Ala Pro Glu Tyr Ala Tyr Lys Ile Val Lys Lys
705                 710                 715                 720
Gly Asp Ser Thr Ile Met Lys Ser Glu
                725
```

<210> 96
<211> 506
<212> PRT
<213> Artificial Sequence

<220>
<223> STF2

<400> 96

```
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
 1               5                  10                  15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
            20                  25                  30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
        35                  40                  45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
    50                  55                  60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65                  70                  75                  80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
                85                  90                  95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
            100                 105                 110
```

```
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
        115                 120                 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
    130                 135                 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145                 150                 155                 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val Gln
                165                 170                 175
Lys Ala Tyr Asp Val Lys Asp Thr Ala Val Thr Thr Lys Ala Tyr Ala
            180                 185                 190
Asn Asn Gly Thr Thr Leu Asp Val Ser Gly Leu Asp Asp Ala Ala Ile
        195                 200                 205
Lys Ala Ala Thr Gly Gly Thr Asn Gly Thr Ala Ser Val Thr Gly Gly
    210                 215                 220
Ala Val Lys Phe Asp Ala Asp Asn Asn Lys Tyr Phe Val Thr Ile Gly
225                 230                 235                 240
Gly Phe Thr Gly Ala Asp Ala Ala Lys Asn Gly Asp Tyr Glu Val Asn
                245                 250                 255
Val Ala Thr Asp Gly Thr Val Thr Leu Ala Ala Gly Ala Thr Lys Thr
            260                 265                 270
Thr Met Pro Ala Gly Ala Thr Thr Lys Thr Glu Val Gln Glu Leu Lys
        275                 280                 285
Asp Thr Pro Ala Val Val Ser Ala Asp Ala Lys Asn Ala Leu Ile Ala
    290                 295                 300
Gly Gly Val Asp Ala Thr Asp Ala Asn Gly Ala Glu Leu Val Lys Met
305                 310                 315                 320
Ser Tyr Thr Asp Lys Asn Gly Lys Thr Ile Glu Gly Gly Tyr Ala Leu
                325                 330                 335
Lys Ala Gly Asp Lys Tyr Tyr Ala Ala Asp Tyr Asp Glu Ala Thr Gly
            340                 345                 350
Ala Ile Lys Ala Lys Thr Thr Ser Tyr Thr Ala Ala Asp Gly Thr Thr
        355                 360                 365
Lys Thr Ala Ala Asn Gln Leu Gly Gly Val Asp Gly Lys Thr Glu Val
    370                 375                 380
Val Thr Ile Asp Gly Lys Thr Tyr Asn Ala Ser Lys Ala Ala Gly His
385                 390                 395                 400
Asp Phe Lys Ala Gln Pro Glu Leu Ala Glu Ala Ala Ala Lys Thr Thr
                405                 410                 415
Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala
            420                 425                 430
Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile
        435                 440                 445
Thr Asn Leu Gly Asn Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg
    450                 455                 460
Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala
465                 470                 475                 480
Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln
                485                 490                 495
Val Pro Gln Asn Val Leu Ser Leu Leu Ala
            500                 505
```

<210> 97
<211> 1614
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2.Blp.wt

<400> 97
atgaaattct tagtcaacgt tgcccttgtt tttatggtcg tatacatttc ttacatctat 60
gcggcccagg ttatcaatac caactccctg tcgttgctca cccaaaataa ccttaataaa 120
agccagagcg cactgggaac cgccatagaa cgcctctcaa gcggcctccg gatcaattct 180
gcaaaagacg acgccgccgg tcaggccatc gcaaaccgct ttaccgccaa tatcaaggga 240
ctgacgcagg cttcgaggaa tgctaacgat ggaataagca tcgctcaaac cacggagggc 300
gccctgaacg agatcaacaa caacctacag cgcgtcaggg agctcgcagt gcagtccgcc 360
aattcgacca actcgcagtc ggacctggac tcgatccaag ccgaaatcac ccagcgcctg 420
aatgagattg accgggtgag cggtcagaca cagtttaacg gcgtgaaggt acttgcacag 480
gataacacac ttacgataca ggtgggcgcc aacgatggtg aaaccataga cattgatctc 540
aaacagatta acagccagac gctcgggttg gatagcctga atgtgcaaaa ggcgtacgac 600
gtgaaagaca cggcggtcac taccaaagcc tacgctaaca atggcactac cttggatgtg 660
agcggattgg atgatgcagc aatcaaggct gctaccggcg gtacgaacgg aaccgcgtcc 720
gtgaccggcg gtgccgtgaa gttcgatgct gacaacaata agtatttcgt caccattgga 780
ggctttactg gcgccgacgc agcaaagaac ggcgactatg aagtgaacgt ggcaaccgat 840
ggaaccgtga cgctggccgc tggtgccacc aagaccacca tgccagccgg cgccacaact 900
aagaccgagg tgcaggagtt aaaggacacc cccgcggtgg ttagcgcaga tgccaaaaac 960
gcgttgatcg ccggcggagt ggatgcaact gatgctaatg gtgcggagct ggttaaaatg 1020
tcgtatacag acaagaatgg taagacgatc gagggcggtt atgcccttaa ggcaggagat 1080
aagtattacg ctgctgatta cgatgaggcg acgggagcta ttaaggccaa gacaacgtca 1140
tacacggcgg cggacggaac gactaagacg gctgccaatc agttgggagg ggttgacggg 1200
aagacagagg tcgttacgat cgatggcaag acatacaacg cctccaaggc cgctggccac 1260
gatttcaaag ctcaacccga actggccgag gccgcggcga aaacaactga gaacccgttg 1320
cagaagattg atgcggccct ggcgcaagta gatgccctgc gctcagacct gggcgccgtt 1380
caaaatcgat tcaattccgc gattacaaac ctgggcaata cagtaaacaa tctatccgag 1440
gccagatccc gcattgaaga ctccgactac gcgacagaag taagtaacat gagtcgtgcc 1500
cagattctgc agcaggccgg cactagtgtc ctggcccagg ccaatcaagt cccgcagaat 1560
gtgctgagcc tactacgaga gttcagtcgt tacccagccc aatggcggcc gctc 1614

<210> 98
<211> 1569
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2.B1p.ng

<400> 98
atgaaattct tagtcaacgt tgcccttgtt tttatggtcg tatacatttc ttacatctat 60
gcggcccaag tgatcaacac caactccctg tccctgctca ctcaaaacaa cctccagaaa 120
tctcagtctg ctctcggtac tgctatcgag cgtctctcct ccggcctgag gatcaactcc 180
gctaaagatg acgccgccgg acaagctatc gctaaccgct ttactgccaa cattaaaggc 240
ctgacccaag cctctaggaa cgccaatgat ggcatttcta tcgctcaaac caccgaaggc 300
gccctgaacg aaatcaataa taacctgcaa cgtgtccgcg aactcgccgt ccagtccgcc 360
caatctacca actctcagtc tgacctggac tctatccaag ctgagatcac ccaaaggctc 420
aacgagatcg atcgtgtgtc tggtcaaact cagtttaacg gcgtcaaagt gctggctcaa 480
gacaacacac tcaccatcca agtaggagcc aatgacggcg agacaatcga tatcgacctg 540
aagcagatca attctcaaac tctgggcctg gactccctga acgtccaaaa ggcttacgac 600
gtgaaggaca ccgctgtgac aaccaaagcc tatgctaatc aaggaacaac cctggacgtg 660
tccggactcg atgacgccgc tatcaaggcc gctactggcg gcactcaggg aactgcctct 720
gtgaccggag gcgctgtgaa gttcgacgcc gataacaaca aatacttcgt cactattggt 780
ggtttcactg gcgctgacgc cgctaagaac ggtgactacg aggtcaacgt cgccactgac 840
ggaacagtga cactggccgc tggcgctacc aagaccacaa tgcctgctgg tgctactact 900
aagacagagg tgcaggaact caaggacacc cctgccgtgg tgtccgccga tgctaagaac 960
gctctcattg ctggtggtgt cgacgctacg gacgccaacg gagccgagct tgtgaaaatg 1020
tcctacaccg acaagaacgg aaagactatc gagggaggtt acgctctgaa ggctggcgat 1080
aagtactacg ccgctgatta tgacgaagct acaggcgcca ttaaagctaa aaccacatct 1140
tataccgctg ccgatggtac taccaagact gctgccaatc agctgggtgg agtcgatgga 1200

aaaaccgaag tggtcacaat cgacggaaag acctatcaag cctccaaggc tgctggccac 1260
gacttcaagg cccaacccga actggccgag gctgctgcta aaaccactga aaacccccctg 1320
cagaaaattg acgctgccct ggcccaagtg gatgccctcc gctccgacct cggcgctgtg 1380
caaaaccgct tcaactccgc tattacaaac ctcggcaata ccgtcaatca gctctctgaa 1440
gctaggtctc gtattgagga ttccgattac gctaccgagg tctcccagat gtcccgtgcc 1500
caaattctcc aacaagccgg cacctccgtc ctcgcccaag ccaatcaggt gccacagaat 1560
gtgctgagc 1569

<210> 99
<211> 63
<212> DNA
<213> Artificial Sequence

<220>
<223> Honey Bee mellitin

<400> 99
atgaaattct tagtcaacgt tgcccttgtt tttatggtcg tatacatttc ttacatctat 60
gcg 63

<210> 100
<211> 3156
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2.HA0s

<400> 100
atgaaattct tagtcaacgt tgcccttgtt tttatggtcg tatacatttc ttacatctat 60
gcggcccagg ttatcaatac caactccctg tcgttgctca cccaaaataa ccttaataaa 120
agccagagcg cactgggaac cgccatagaa cgcctctcaa gcggcctccg gatcaattct 180
gcaaaagacg acgccgccgg tcaggccatc gcaaaccgct ttaccgccaa tatcaaggga 240
ctgacgcagg cttcgaggaa tgctaacgat ggaataagca tcgctcaaac cacggagggc 300
gccctgaacg agatcaacaa caacctacag cgcgtcaggg agctcgcagt gcagtccgcc 360
aattcgacca actcgcagtc ggacctggac tcgatccaag ccgaaatcac ccagcgcctg 420
aatgagattg accgggtgag cggtcagaca cagtttaacg gcgtgaaggt acttgcacag 480
gataacacac ttacgataca ggtgggcgcc aacgatggtg aaaccataga cattgatctc 540
aaacagatta acagccagac gctcgggttg gatagcctga atgtgcaaaa ggcgtacgac 600
gtgaaagaca cggcggtcac taccaaagcc tacgctaaca atggcactac cttggatgtg 660
agcggattgg atgatgcagc aatcaaggct gctaccggcg gtacgaacgg aaccgcgtcc 720
gtgaccggcg gtgccgtgaa gttcgatgct gacaacaata agtatttcgt caccattgga 780
ggctttactg gcgccgacgc agcaaagaac ggcgactatg aagtgaacgt ggcaaccgat 840
ggaaccgtga cgctggccgc tggtgccacc aagaccacca tgccagccgg cgccacaact 900
aagaccgagg tgcaggagtt aaaggacacc cccgcggtgg ttagcgcaga tgccaaaaac 960
gccttgatcg ccggcggagt ggatgcaact gatgctaatg gtgcggagct ggttaaaatg 1020
tcgtatacag acaagaatgg taagacgatc gagggcggtt atgcccttaa ggcaggagat 1080
aagtattacg ctgctgatta cgatgaggcg acgggagcta ttaaggccaa gacaacgtca 1140
tacacggcgg cggacggaac gactaagacg gctgccaatc agttgggagg ggttgacggg 1200
aagacagagg tcgttacgat cgatggcaag acatacaacg cttccaaggc cgctggccac 1260
gatttcaaag ctcaacccga actggccgag gccgcggcga aaacaactga gaacccgttg 1320
cagaagattg atgcggccct ggcgcaagta gatgccctgc gctcagacct gggcgccgtt 1380
caaaatcgat tcaattccgc gattacaaac ctgggcaata cagtaaacaa tctatccgag 1440
gccagatccc gcattgaaga ctccgactac gcgacagaag taagtaacat gagtcgtgcc 1500
cagattctgc agcaggccgg cactagtgtc ctggcccagg ccaatcaagt cccgcagaat 1560
gtgctgagcc tactagcatc aggttccggc tcaggttccg acaccatttg cattggatac 1620
catgcaaaca actcaaccga tactgttgat accgtccttg agaagaacgt taccgtcacg 1680
cactcggtca acctattaga ggatagccac aacggaaagc tgtgccgtct gaaaggcata 1740
gcgccactgc agctcggcaa atgtaacatc gcaggttggc tccttggtaa cccggagtgc 1800

gacccccctcc tccctgtacg atcctggagt tatatcgtgg agacccccaa tagcgagaac 1860
ggaatttgct acccaggaga tttcatagac tacgaggagt tgcgcgagca gctttcgtct 1920
gtgagcagct tcgaaaggtt cgagatattc ccgaaggaga gcagctggcc gaatcataac 1980
actaacggtg tgacagccgc ctgcagtcat gaaggaaaga gttcattcta tcgcaacctg 2040
ctgtggttga cggagaaaga gggcagctac cctaagttga agaactccta tgtgaacaaa 2100
aaaggcaagg aggttctggt gctgtggggc atacaccacc cccccaatag caaggagcag 2160
cagaatctgt accaaaacga gaatgcctat gtgagcgtgg tcactagtaa ctataaccgt 2220
cggttcactc ccgagatcgc cgagcgtccg aaggtgaggg accaggcagg ccggatgaac 2280
tactactgga ccctattgaa gccaggggac acgattatct tcgaggcaaa cggaaacctc 2340
atagcgccga tgtacgcctt cgccctgagc cgtggctttg gatcggggat catcacgtct 2400
aacgcctcga tgcacgaatg taataccaaa tgccagaccc cactgggtgc tatcaactcg 2460
tccttaccct atcaaaatat acatccggtc accataggcg agtgtcccaa atatgtcaga 2520
tccgccaagt tgcggatggt gaccggcctc cgcaatattc ctagtattca gtcacgcggc 2580
ttgttcggcg ccatcgctgg tttcatcgaa ggcgggtgga caggcatgat tgatggctgg 2640
tatggctatc accaccagaa cgagcagggc tcgggctacg cggctgacca gaagtcgact 2700
cagaatgcca tcaatggcat cacgaataag gtgaacacgg tcattgaaaa gatgaacatt 2760
caatttacag ccgtaggaaa agagtttaat aaactggaaa aaagaatgga gaatctgaat 2820
aagaaggtgg acgacggatt tttggacatc tggacgtaca acgccgagct gctggttctg 2880
ctggaaaatg agcgaacact ggattttcat gattctaacg taaagaatct gtacgagaag 2940
gtgaagtccc aactaaagaa taatgccaag gaaatcggaa atggatgctt tgagttttac 3000
cacaagtgcg ataatgagtg catggaatcc gtgcgaaatg gtacatacga ttacccaaag 3060
tactccgaag aatccaagct aaatcgcgaa aaggttgatg gtgttaaact tgaatccatg 3120
ggtatttacc aacaccatca tcaccaccat taatag 3156

<210> 101
<211> 67
<212> DNA
<213> Artificial Sequence

<220>
<223> FOR PRIMER

<400> 101
agtcgctgag cctactagca tcaggttccg gctcaggttc cgacaccatt tgcattggat 60
accatgc 67

<210> 102
<211> 55
<212> DNA
<213> Artificial Sequence

<220>
<223> REV PRIMER

<400> 102
agtcgcatgc ctattaatgg tggtgatgat ggtgttggta aatacccatg gattc 55

<210> 103
<211> 1557
<212> DNA
<213> Artificial Sequence

<220>
<223> HA0s PR8

<400> 103
gacaccattt gcattggata ccatgcaaac aactcaaccg atactgttga taccgtcctt 60
gagaagaacg ttaccgtcac gcactcggtc aacctattag aggatagcca caacggaaag 120
ctgtgccgtc tgaaaggcat agcgccactg cagctcggca aatgtaacat cgcaggttgg 180

ctccttggta acccggagtg cgacccctc ctccctgtac gatcctggag ttatatcgtg 240
gagacccca atagcgagaa cggaatttgc tacccaggag atttcataga ctacgaggag 300
ttgcgcgagc agctttcgtc tgtgagcagc ttcgaaaggt tcgagatatt cccgaaggag 360
agcagctggc cgaatcataa cactaacggt gtgacagccg cctgcagtca tgaaggaaag 420
agttcattct atcgcaacct gctgtggttg acggagaaag agggcagcta ccctaagttg 480
aagaactcct atgtgaacaa aaaaggcaag gaggttctgg tgctgtgggg catacaccac 540
ccccccaata gcaaggagca gcagaatctg taccaaaacg agaatgccta tgtgagcgtg 600
gtcactagta actataaccg tcggttcact cccgagatcg ccgagcgtcc gaaggtgagg 660
gaccaggcag gccggatgaa ctactactgg accctattga agccagggga cacgattatc 720
ttcgaggcaa acggaaacct catagcgccg atgtacgcct tcgccctgag ccgtggcttt 780
ggatcgggga tcatcacgtc taacgcctcg atgcacgaat gtaataccaa atgccagacc 840
ccactgggtg ctatcaactc gtccttaccc tatcaaaata tacatccggt caccataggc 900
gagtgtccca aatatgtcag atccgccaag ttgcggatgg tgaccggcct ccgcaatatt 960
cctagtattc agtcacgcgg cttgttcggc gccatcgctg gtttcatcga aggcgggtgg 1020
acaggcatga ttgatggctg gtatggctat caccaccaga acgagcaggg ctcgggctac 1080
gcggctgacc agaagtcgac tcagaatgcc atcaatggca tcacgaataa ggtgaacacg 1140
gtcattgaaa agatgaacat tcaatttaca gccgtaggaa aagagtttaa taaactggaa 1200
aaaagaatgg agaatctgaa taagaaggtg gacgacggat ttttggacat ctggacgtac 1260
aacgccgagc tgctggttct gctggaaaat gagcgaacac tggattttca tgattctaac 1320
gtaaagaatc tgtacgagaa ggtgaagtcc caactaaaga ataatgccaa ggaaatcgga 1380
aatggatgct ttgagtttta ccacaagtgc gataatgagt gcatggaatc cgtgcgaaat 1440
ggtacatacg attacccaaa gtactccgaa gaatccaagc taaatcgcga aaaggttgat 1500
ggtgttaaac ttgaatccat gggtatttac caacaccatc atcaccacca ttaatag 1557

<210> 104
<211> 2439
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2.HA1-1

<400> 104
atgaaattct tagtcaacgt tgcccttgtt tttatggtcg tatacatttc ttacatctat 60
gcggcccagg ttatcaatac caactccctg tcgttgctca cccaaaataa ccttaataaa 120
agccagagcg cactgggaac cgccatagaa cgcctctcaa gcggcctccg gatcaattct 180
gcaaaagacg acgccgccgg tcaggccatc gcaaaccgct ttaccgccaa tatcaaggga 240
ctgacgcagg cttcgaggaa tgctaacgat ggaataagca tcgctcaaac cacggagggc 300
gccctgaacg agatcaacaa caacctacag cgcgtcaggg agctcgcagt gcagtccgcc 360
aattcgacca actcgcagtc ggacctggac tcgatccaag ccgaaatcac ccagcgcctg 420
aatgagattg accgggtgag cggtcagaca cagtttaacg gcgtgaaggt acttgcacag 480
gataacacac ttacgataca ggtgggcgcc aacgatggtg aaaccataga cattgatctc 540
aaacagatta acagccagac gctcgggttg gatagcctga atgtgcaaaa ggcgtacgac 600
gtgaaagaca cggcggtcac taccaaagcc tacgctaaca atggcactac cttggatgtg 660
agcggattgg atgatgcagc aatcaaggct gctaccggcg gtacgaacgg aaccgcgtcc 720
gtgaccggcg gtgccgtgaa gttcgatgct gacaacaata agtatttcgt caccattgga 780
ggctttactg gcgccgacgc agcaaagaac ggcgactatg aagtgaacgt ggcaaccgat 840
ggaaccgtga cgctggccgc tggtgccacc aagaccacca tgccagccgg cgccacaact 900
aagaccgagg tgcaggagtt aaaggacacc cccgcggtgg ttagcgcaga tgccaaaaac 960
gccttgatcg ccggcggagt ggatgcaact gatgctaatg gtgcggagct ggttaaaatg 1020
tcgtatacag acaagaatgg taagacgatc gagggcggtt atgcccttaa ggcaggagat 1080
aagtattacg ctgctgatta cgatgaggcg acgggagcta ttaaggccaa gacaacgtca 1140
tacacggcgg cggacggaac gactaagacg gctgccaatc agttgggagg ggttgacggg 1200
aagacagagg tcgttacgat cgatggcaag acatacaacg cttccaaggc cgctggccac 1260
gatttcaaag ctcaacccga actggccgag gccgcggcga aaacaactga gaacccgttg 1320
cagaagattg atgcggccct ggcgcaagta gatgccctgc gctcagacct gggcgccgtt 1380
caaaatcgat tcaattccgc gattacaaac ctgggcaata cagtaaacaa tctatccgag 1440
gccagatccc gcattgaaga ctccgactac gcgacagaag taagtaacat gagtcgtgcc 1500
cagattctgc agcaggccgg cactagtgtc ctggcccagg ccaatcaagt cccgcagaat 1560

gtgctgagcc tactagcatc aggttccggc tcaggttcca gccacaacgg aaagctgtgc 1620
cgtctgaaag gcatagcgcc actgcagctc ggcaaatgta acatcgcagg ttggctcctt 1680
ggtaacccgg agtgcgaccc cctcctccct gtacgatcct ggagttatat cgtggagacc 1740
cccaatagcg agaacggaat ttgctaccca ggagatttca tagactacga ggagttgcgc 1800
gagcagcttt cgtctgtgag cagcttcgaa aggttcgaga tattcccgaa ggagagcagc 1860
tggccgaatc ataacactaa cggtgtgaca gccgcctgca gtcatgaagg aaagagttca 1920
ttctatcgca acctgctgtg gttgacggag aaagagggca gctaccctaa gttgaagaac 1980
tcctatgtga acaaaaaagg caaggaggtt ctggtgctgt ggggcataca ccacccccc 2040
aatagcaagg agcagcagaa tctgtaccaa aacgagaatg cctatgtgag cgtggtcact 2100
agtaactata accgtcggtt cactcccgag atcgccgagc gtccgaaggt gagggaccag 2160
gcaggccgga tgaactacta ctggacccta ttgaagccag gggacacgat tatcttcgag 2220
gcaaacggaa acctcatagc gccgatgtac gccttcgccc tgagccgtgg ctttggatcg 2280
gggatcatca cgtctaacgc ctcgatgcac gaatgtaata ccaaatgcca gaccccactg 2340
ggtgctatca actcgtcctt accctatcaa aatatacatc cggtcaccat aggcgagtgt 2400
cccaaatatg tcagacacca tcatcaccac cattaatag 2439

<210> 105
<211> 66
<212> DNA
<213> Artificial Sequence

<220>
<223> FOR PRIMER

<400> 105
agtcgctgag cctactagca tcaggttccg gctcaggttc cagccacaac ggaaagctgt 60
gccgtc 66

<210> 106
<211> 67
<212> DNA
<213> Artificial Sequence

<220>
<223> REV PRIMER

<400> 106
agtcgcatgc ctattaatgg tggtgatgat ggtgtctgac atatttggga cactcgccta 60
tggtgac 67

<210> 107
<211> 2292
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2.HA1-2

<400> 107
atgaaattct tagtcaacgt tgcccttgtt tttatggtcg tatacatttc ttacatctat 60
gcggcccagg ttatcaatac caactccctg tcgttgctca cccaaaataa ccttaataaa 120
agccagagcg cactgggaac cgccatagaa cgcctctcaa gcggcctccg gatcaattct 180
gcaaaagacg acgccgccgg tcaggccatc gcaaaccgct ttaccgccaa tatcaaggga 240
ctgacgcagg cttcgaggaa tgctaacgat gggataagca tcgctcaaac cacggagggc 300
gccctgaacg agatcaacaa caacctacag cgcgtcaggg agctcgcagt gcagtccgcc 360
aattcgacca actcgcagtc ggacctggac tcgatccaag ccgaaatcac ccagcgcctg 420
aacgagattg accgggtgag cggtcagaca cagtttaacg gcgtgaaggt acttgcgcag 480
gataacacac ttacgataca ggtgggcgcc aacgatggtg aaaccataga cattgatctc 540
aaacagatta acagccagac gctcgggttg gatagcctga atgtgcaaaa ggcgtacgac 600

gtgaaagaca cggcggtcac taccaaagcc tacgctaaca atggcactac cttggatgtg 660
agcggattgg atgatgcagc aatcaaggct gctaccggcg gtacgaacgg aaccgcgtcc 720
gtgaccggcg gtgccgtgaa gttcgatgct gacaacaata agtatttcgt caccattgga 780
ggctttactg gcgccgacgc agcaaagaac ggcgactatg aagtgaacgt ggcaaccgat 840
ggaaccgtga cgctggccgc tggtgccacc aagaccacca tgccagccgg cgccacaact 900
aagaccgagg tgcaggagtt aaaggacacc cccgcggtgg ttagcgcaga tgccaaaaac 960
gcgttgatcg ccggcggagt ggatgcaact gatgctaatg gtgcggagct ggttaaaatg 1020
tcgtatacag acaagaatgg taagacgatc gagggcggtt atgcccttaa ggcaggagat 1080
aagtattacg ctgctgatta cgatgaggcg acgggagcta ttaaggccaa gacaacgtca 1140
tacacggcgg cggacggaac gactaagacg gctgccaatc agttgggagg ggttgacggg 1200
aagacagagg tcgttacgat cgatggcaag acatacaacg cctccaaggc cgctggccac 1260
gatttcaaag ctcaacccga actggccgag gccgcggcga aaacaactga gaacccgttg 1320
cagaagattg atgcggccct ggcgcaagta gatgccctgc gctcagacct gggcgccgtt 1380
caaaatcgat tcaattccgc gattacaaac ctgggcaata cagtaaacaa tctatccgag 1440
gccagatccc gcattgaaga ctccgactac gcgacagaag taagtaacat gagtcgtgcc 1500
cagattctgc agcaggccgg cactagtgtc ctggcccagg ccaatcaagt cccgcagaat 1560
gtgctgagcc tactagcatc aggttccggc tcaggttcca aaggcatagc gccactgcag 1620
ctcggcaaat gtaacatcgc aggttggctc cttggtaacc cggagtgcga ccccctcctc 1680
cctgtacgat cctggagtta tatcgtggag acccccaata gcgagaacgg aatttgctac 1740
ccaggagatt tcatagacta cgaggagttg cgcgagcagc tttcgtctgt gagcagcttc 1800
gagaggttcg aaatcttccc gaaggagagc agctggccga atcataacac taacggtgtg 1860
acagccgcct gcagtcatga aggaaagagt tcattctatc gcaacctgct gtggttgacg 1920
gagaaagagg gcagctaccc taagttgaag aactcctatg tgaacaaaaa aggcaaggag 1980
gttctggtgc tgtggggcat acaccacccc cccaatagca aggagcagca gaatctgtac 2040
caaaacgaga atgcctatgt gagcgtggtc actagtaact ataaccgtcg gttcactccc 2100
gagatcgccg agcgtccgaa ggtgagggac caggcaggcc ggatgaacta ctactggacc 2160
ctattgaagc caggggacac gattatcttc gaggcaaacg gaaacctcat agcgccgatg 2220
tacgcgttcg ccctgagccg tggctttgga tcggggatca tcacgtctca ccatcatcac 2280
caccattaat ag 2292

<210> 108
<211> 66
<212> DNA
<213> Artificial Sequence

<220>
<223> FOR PRIMER

<400> 108
agtcgctgag cctactagca tcaggttccg gctcaggttc caaaggcata gcgccactgc 60
agctcg 66

<210> 109
<211> 70
<212> DNA
<213> Artificial Sequence

<220>
<223> REV PRIMER

<400> 109
agtcgcatgc ctattaatgg tggtgatgat ggtgagacgt gatgatcccc gatccaaagc 60
cacggctcag 70

<210> 110
<211> 2292
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2.HA1-2mut

<400> 110
atgaaattct tagtcaacgt tgcccttgtt tttatggtcg tatacatttc ttacatctat 60
gcggcccagg ttatcaatac caactccctg tcgttgctca cccaaaataa ccttaataaa 120
agccagagcg cactgggaac cgccatagaa cgcctctcaa gcggcctccg gatcaattct 180
gcaaaagacg acgccgccgg tcaggccatc gcaaaccgct ttaccgccaa tatcaaggga 240
ctgacgcagg cttcgaggaa tgctaacgat ggaataagca tcgctcaaac cacggagggc 300
gccctgaacg agatcaacaa caacctacag cgcgtcaggg agctcgcagt gcagtccgcc 360
aattcgacca actcgcagtc ggacctggac tcgatccaag ccgaaatcac ccagcgcctg 420
aatgagattg accgggtgag cggtcagaca cagtttaacg gcgtgaaggt acttgcacag 480
gataacacac ttacgataca ggtgggcgcc aacgatggtg aaaccataga cattgatctc 540
aaacagatta acagccagac gctcgggttg gatagcctga atgtgcaaaa ggcgtacgac 600
gtgaaagaca cggcggtcac taccaaagcc tacgctaaca atggcactac cttggatgtg 660
agcggattgg atgatgcagc aatcaaggct gctaccggcg gtacgaacgg aaccgcgtcc 720
gtgaccggcg gtgccgtgaa gttcgatgct gacaacaata agtatttcgt caccattgga 780
ggctttactg gcgccgacgc agcaaagaac ggcgactatg aagtgaacgt ggcaaccgat 840
ggaaccgtga cgctggccgc tggtgccacc aagaccacca tgccagccgg cgccacaact 900
aagaccgagg tgcaggagtt aaaggacacc cccgcggtgg ttagcgcaga tgccaaaaac 960
gcgttgatcg ccggcggagt ggatgcaact gatgctaatg gtgcggagct ggttaaaatg 1020
tcgtatacag acaagaatgg taagacgatc gagggcggtt atgcccttaa ggcaggagat 1080
aagtattacg ctgctgatta cgatgaggcg acgggagcta ttaaggccaa gacaacgtca 1140
tacacggcgg cggacggaac gactaagacg gctgccaatc agttgggagg ggttgacggg 1200
aagacagagg tcgttacgat cgatggcaag acatacaacg cctccaaggc cgctggccac 1260
gatttcaaag ctcaacccga actggccgag gccgcggcga aaacaactga gaacccgttg 1320
cagaagattg atgcggccct ggcgcaagta gatgccctgc gctcagacct gggcgccgtt 1380
caaaatcgat tcaattccgc gattacaaac ctgggcaata cagtaaacaa tctatccgag 1440
gccagatccc gcattgaaga ctccgactac gcgacagaag taagtaacat gagtcgtgcc 1500
cagattctgc agcaggccgg cactagtgtc ctggcccagg ccaatcaagt cccgcagaat 1560
gtgctgagcc tgctggcttc tggatctgga tctggttcta aaggcgctgc ccctctgcaa 1620
ctcggcaagt gcaatattgc ggggtggctg ttgggcaacc cagaatgtga ccccctcctg 1680
cccgtccgtt cttggtctga catcgctgaa acacctaact ccgagaacgg catctgttac 1740
ccgggcgact tcattgacta cgaggagctc cgcgaacagc tgtcttctgt ctcttctttc 1800
gaacgtttcg aaattttccc aaaggaatcc tcctggccaa accataacac caacggagtg 1860
accgccgcct gtagccatga gggcaagtct tctttctacc gtaatctgct gtggctgact 1920
gaaaaagagg gttcttatcc caaactgaag aactcttatg tcaacaagaa gggcaaagag 1980
gtcctggtgc tgtggggtat ccaccacccc cccaactcca aggagcagca gaatctgtac 2040
caaaacgaaa atgcttacgt gtctgtcgtg acatctaact acaaccggcg cttcacgccc 2100
gaaatcgccg agcgtcccaa agtgcgcgat caggccggaa ggatgaacta ctactggacc 2160
ctgctcaaac ccggagatac catcatattc gaagccaatg gcaatctcat cgctcccatg 2220
tatgctttcg ctctctctcg cggattcggt tccggcataa tcactagcca ccaccaccac 2280
caccactaat ag 2292

<210> 111
<211> 692
<212> DNA
<213> Artificial Sequence

<220>
<223> HA1-2mut

<400> 111
aaggcgctgc ccctctgcaa ctcggcaagt gcaatattgc ggggtggctg ttgggcaacc 60
cagaatgtga ccccctcctg cccgtccgtt cttggtctga catcgctgaa acacctaact 120
ccgagaacgg catctgttac ccgggcgact tcattgacta cgaggagctc cgcgaacagc 180
tgtcttctgt ctcttctttc gaacgtttcg aaattttccc aaaggaatcc tcctggccaa 240
accataacac caacggagtg accgccgcct gtagccatga gggcaagtct tctttctacc 300
gtaatctgct gtggctgact gaaaaagagg gttcttatcc caaactgaag aactcttatg 360

tcaacaagaa gggcaaagag gtcctggtgc tgtggggtat ccaccacccc cccaactcca 420
aggagcagca gaatctgtac caaaacgaaa atgcttacgt gtctgtcgtg acatctaact 480
acaaccggcg cttcacgccc gaaatcgccg agcgtcccaa agtgcgcgat caggccggaa 540
ggatgaacta ctactggacc ctgctcaaac ccggagatac catcatattc gaagccaatg 600
gcaatctcat cgctcccatg tatgctttcg ctctctctcg cggattcggt tccggcataa 660
tcactagcca ccaccaccac caccactaat ag 692

<210> 112
<211> 2151
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2.HA1-3

<400> 112
atgaaattct tagtcaacgt tgcccttgtt tttatggtcg tatacatttc ttacatctat 60
gcggcccagg ttatcaatac caactccctg tcgttgctca cccaaaataa ccttaataaa 120
agccagagcg cactgggaac cgccatagaa cgcctctcaa gcggcctccg gatcaattct 180
gcaaaagacg acgccgccgg tcaggccatc gcaaaccgct ttaccgccaa tatcaaggga 240
ctgacgcagg cttcgaggaa tgctaacgat gggataagca tcgctcaaac cacggagggc 300
gccctgaacg agatcaacaa caacctacag cgcgtcaggg agctcgcagt gcagtccgcc 360
aattcgacca actcgcagtc ggacctggac tcgatccaag ccgaaatcac ccagcgcctg 420
aacgagattg accgggtgag cggtcagaca cagtttaacg gcgtgaaggt acttgcgcag 480
gataacacac ttacgataca ggtgggcgcc aacgatggtg aaaccataga cattgatctc 540
aaacagatta acagccagac gctcgggttg gatagcctga atgtgcaaaa ggcgtacgac 600
gtgaaagaca cggcggtcac taccaaagcc tacgctaaca atggcactac cttggatgtg 660
agcggattgg atgatgcagc aatcaaggct gctaccggcg gtacgaacgg aaccgcgtcc 720
gtgaccggcg gtgccgtgaa gttcgatgct gacaacaata agtatttcgt caccattgga 780
ggctttactg gcgccgacgc agcaaagaac ggcgactatg aagtgaacgt ggcaaccgat 840
ggaaccgtga cgctggccgc tggtgccacc aagaccacca tgccagccgg cgccacaact 900
aagaccgagg tgcaggagtt aaaggacacc cccgcggtgg ttagcgcaga tgccaaaaac 960
gcgttgatcg ccggcggagt ggatgcaact gatgctaatg gtgcggagct ggttaaaatg 1020
tcgtatacag acaagaatgg taagacgatc gagggcggtt atgcccttaa ggcaggagat 1080
aagtattacg ctgctgatta cgatgaggcg acgggagcta ttaaggccaa gacaacgtca 1140
tacacggcgg cggacggaac gactaagacg gctgccaatc agttgggagg ggttgacggg 1200
aagacagagg tcgttacgat cgatggcaag acatacaacg cctccaaggc cgctggccac 1260
gatttcaaag ctcaacccga actggccgag gccgcggcga aaacaactga gaacccgttg 1320
cagaagattg atgcggccct ggcgcaagta gatgccctgc gctcagacct gggcgccgtt 1380
caaaatcgat tcaattccgc gattacaaac ctgggcaata cagtaaacaa tctatccgag 1440
gccagatccc gcattgaaga ctccgactac gcgacagaag taagtaacat gagtcgtgcc 1500
cagattctgc agcaggccgg cactagtgtc ctggcccagg ccaatcaagt cccgcagaat 1560
gtgctgagcc tactagcatc aggttccggc tcaggttcca atagcgagaa cggaatttgc 1620
tacccaggag atttcataga ctacgaggag ttgcgcgagc agctttcgtc tgtgagcagc 1680
ttcgagaggt tcgaaatctt cccgaaggag agcagctggc cgaatcataa cactaacggt 1740
gtgacagccg cctgcagtca tgaaggaaag agttcattct atcgcaacct gctgtggttg 1800
acggagaagg agggcagcta ccctaagctg aagaactcct atgtgaacaa aaaaggcaag 1860
gaggttctgg tgctgtgggg catacaccac ccccccaata gcaaggagca gcagaatctg 1920
taccaaaacg agaatgccta tgtgagcgtg gtcactagta actataaccg tcggttcact 1980
cccgagatcg ccgagcgtcc gaaggtgagg gaccaggcag gccggatgaa ctactactgg 2040
accctattga agccagggga cacgattatc ttcgaggcaa acggaaacct catagcgccg 2100
atgtacgcct tcgccctgag ccgtggccac catcatcacc accattaata g 2151

<210> 113
<211> 66
<212> DNA
<213> Artificial Sequence

<220>

<223> FOR PRIMER

<400> 113
agtcgctgag cctactagca tcaggttccg gctcaggttc caatagcgag aacggaattt 60
gctacc 66

<210> 114
<211> 54
<212> DNA
<213> Artificial Sequence

<220>
<223> REV PRIMER

<400> 114
agtcgcatgc ctattaatgg tggtgatgat ggtggccacg gctcagggcg aagg 54

<210> 115
<211> 2151
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2.HA1-3mut

<400> 115
atgaaattct tagtcaacgt tgcccttgtt tttatggtcg tatacatttc ttacatctat 60
gcggcccagg ttatcaatac caactccctg tcgttgctca cccaaaataa ccttaataaa 120
agccagagcg cactgggaac cgccatagaa cgcctctcaa gcggcctccg gatcaattct 180
gcaaaagacg acgccgccgg tcaggccatc gcaaaccgct ttaccgccaa tatcaaggga 240
ctgacgcagg cttcgaggaa tgctaacgat ggaataagca tcgctcaaac cacggagggc 300
gccctgaacg agatcaacaa caacctacag cgcgtcaggg agctcgcagt gcagtccgcc 360
aattcgacca actcgcagtc ggacctggac tcgatccaag ccgaaatcac ccagcgcctg 420
aatgagattg accgggtgag cggtcagaca cagtttaacg gcgtgaaggt acttgcacag 480
gataacacac ttacgataca ggtgggcgcc aacgatggtg aaaccataga cattgatctc 540
aaacagatta acagccagac gctcgggttg gatagcctga atgtgcaaaa ggcgtacgac 600
gtgaaagaca cggcggtcac taccaaagcc tacgctaaca atggcactac cttggatgtg 660
agcggattgg atgatgcagc aatcaaggct gctaccggcg gtacgaacgg aaccgcgtcc 720
gtgaccggcg gtgccgtgaa gttcgatgct gacaacaata agtatttcgt caccattgga 780
ggctttactg gcgccgacgc agcaaagaac ggcgactatg aagtgaacgt ggcaaccgat 840
ggaaccgtga cgctggccgc tggtgccacc aagaccacca tgccagccgg cgccacaact 900
aagaccgagg tgcaggagtt aaaggacacc cccgcggtgg ttagcgcaga tgccaaaaac 960
gcgttgatcg ccggcggagt ggatgcaact gatgctaatg gtgcggagct ggttaaaatg 1020
tcgtatacag acaagaatgg taagacgatc gagggcggtt atgcccttaa ggcaggagat 1080
aagtattacg ctgctgatta cgatgaggcg acgggagcta ttaaggccaa gacaacgtca 1140
tacacggcgg cggacggaac gactaagacg gctgccaatc agttgggagg ggttgacggg 1200
aagacagagg tcgttacgat cgatggcaag acatacaacg cctccaaggc cgctggccac 1260
gatttcaaag ctcaacccga actggccgag gccgcggcga aaacaactga gaacccgttg 1320
cagaagattg atgcggccct ggcgcaagta gatgccctgc gctcagacct gggcgccgtt 1380
caaaatcgat tcaattccgc gattacaaac ctgggcaata cagtaaacaa tctatccgag 1440
gccagatccc gcattgaaga ctccgactac gcgacagaag taagtaacat gagtcgtgcc 1500
cagattctgc agcaggccgg cactagtgtc ctggcccagg ccaatcaagt cccgcagaat 1560
gtgctgagcc tgctggcttc tggatctgga tctggttcta actccgagaa cgaaatctgt 1620
tacccgggcg acttcattga caaagaggag ctccgcgaac agctgtcttc tgtctcttct 1680
ttcgaacgtt tcgaaatttt cccaaaggaa tcctcctggc caaaccataa caccaacgga 1740
gtgaccgccg cctgtagcca tgagggcaag tcttctttct accgtaatct gctgtggctg 1800
actgaaaaag agggttctta tcccaaactg aagaactctt atgtcaacaa gaagggcaaa 1860
gaggtcctgg tgctgtgggg tatccaccac ccccccaact ccaaggagca gcagaatctg 1920
taccaaaacg aaaatgctta cgtgtctgtc gtgacatcta actacaaccg gcgcttcacg 1980

cccgaaatcg ccgagcgtcc caaagtgcgc gatcaggccg gaaggatgaa ctactactgg 2040
accctgctca aacccggaga taccatcata ttcgaagcca atggcaatct catcgctccc 2100
atgtatgctg ctgctctctc tcgcggacac caccaccacc accactaata g 2151

<210> 116
<211> 552
<212> DNA
<213> Artificial Sequence

<220>
<223> HA1-3mut PR8

<400> 116
aactccgaga acgaaatctg ttacccgggc gacttcattg acaaagagga gctccgcgaa 60
cagctgtctt ctgtctcttc tttcgaacgt ttcgaaattt tcccaaagga atcctcctgg 120
ccaaaccata acaccaacgg agtgaccgcc gcctgtagcc atgagggcaa gtcttctttc 180
taccgtaatc tgctgtggct gactgaaaaa gagggttctt atcccaaact gaagaactct 240
tatgtcaaca agaagggcaa agaggtcctg gtgctgtggg gtatccacca ccccccaac 300
tccaaggagc agcagaatct gtaccaaaac gaaaatgctt acgtgtctgt cgtgacatct 360
aactacaacc ggcgcttcac gcccgaaatc gccgagcgtc ccaaagtgcg cgatcaggcc 420
ggaaggatga actactactg gaccctgctc aaacccggag ataccatcat attcgaagcc 480
aatggcaatc tcatcgctcc catgtatgct gctgctctct ctcgcggaca ccaccaccac 540
caccactaat ag 552

<210> 117
<211> 3156
<212> DNA
<213> Artificial Sequence

<220>
<223> ngSTF2.HA0s

<400> 117
atgaaattct tagtcaacgt tgcccttgtt tttatggtcg tatacatttc ttacatctat 60
gcggcccaag tgatcaacac caactccctg tccctgctca ctcaaaacaa cctccagaaa 120
tctcagtctg ctctcggtac tgctatcgag cgtctctcct ccggcctgag gatcaactcc 180
gctaaagatg acgccgccgg acaagctatc gctaaccgct ttactgccaa cattaaaggc 240
ctgacccaag cctctaggaa cgccaatgat ggcatttcta tcgctcaaac caccgaaggc 300
gccctgaacg aaatcaataa taacctgcaa cgtgtccgcg aactcgccgt ccagtccgcc 360
caatctacca actctcagtc tgacctggac tctatccaag ctgagatcac ccaaaggctc 420
aacgagatcg atcgtgtgtc tggtcaaact cagtttaacg gcgtcaaagt gctggctcaa 480
gacaacacac tcaccatcca agtaggagcc aatgacggcg agacaatcga tatcgacctg 540
aagcagatca attctcaaac tctgggcctg gactccctga acgtccaaaa ggcttacgac 600
gtgaaggaca ccgctgtgac aaccaaagcc tatgctaatc aaggaacaac cctggacgtg 660
tccggactcg atgacgccgc tatcaaggcc gctactggcg gcactcaggg aactgcctct 720
gtgaccggag gcgctgtgaa gttcgacgcc gataacaaca aatacttcgt cactattggt 780
ggtttcactg gcgctgacgc cgctaagaac ggtgactacg aggtcaacgt cgccactgac 840
ggaacagtga cactggccgc tggcgctacc aagaccacaa tgcctgctgg tgctactact 900
aagacagagg tgcaggaact caaggacacc cctgccgtgg tgtccgccga tgctaagaac 960
gctctcattg ctggtggtgt cgacgctacg gacgccaacg gagccgagct tgtgaaaatg 1020
tcctacaccg acaagaacgg aaagactatc gagggaggtt acgctctgaa ggctggcgat 1080
aagtactacg ccgctgatta tgacgaagct acaggcgcca ttaaagctaa aaccacatct 1140
tataccgctg ccgatggtac taccaagact gctgccaatc agctgggtgg agtcgatgga 1200
aaaaccgaag tggtcacaat cgacggaaag acctatcaag cctccaaggc tgctggccac 1260
gacttcaagg cccaacccga actggccgag gctgctgcta aaaccactga aaacccctg 1320
cagaaaattg acgctgccct ggcccaagtg gatgccctcc gctccgacct cggcgctgtg 1380
caaaaccgct tcaactccgc tattacaaac ctcggcaata ccgtcaatca actctctgaa 1440
gctaggtctc gtattgagga ttccgattac gctaccgagg tctcccagat gtcccgtgcc 1500
caaattctcc aacaagccgg cacctccgtc ctcgcccaag ccaatcaggt gccacagaat 1560

gtgctgagcc tgctggcttc tggatctgga tctggttctg acaccatctg cattggctac 1620
catgcccagc agtctacaga tacagtcgat accgtcctgg agaaacaagt caccgtcacg 1680
cactccgtga acctcctgga ggactcccat aatggcaaat tgtgtcggct gaaaggcatc 1740
gcccctctgc aactcggcaa gtgcaatatt gcggggtggc tgttgggcaa cccagaatgt 1800
gacccccctcc tgcccgtccg ttcttggtct tacatcgtgg aaacacctaa ctccgagaac 1860
ggcatctgtt acccgggcga cttcattgac tacgaggagc tccgcgaaca gctgtcttct 1920
gtctcttctt tcgaacgttt cgaaattttc ccaaaggaat cctcctggcc aaaccataac 1980
accaacggag tgaccgccgc ctgtagccat gagggcaagt cttctttcta ccgtaatctg 2040
ctgtggctga ctgaaaaaga gggttcttat cccaaactga agaactctta tgtcaacaag 2100
aagggcaaag aggtcctggt gctgtggggt atccaccacc cccccaactc caaggagcag 2160
cagaatctgt accaaaacga aaatgcttac gtgtctgtcg tgacatctaa ctacaaccgg 2220
cgcttcacgc ccgaaatcgc cgagcgtccc aaagtgcgcg atcaggccgg aaggatgaac 2280
tactactgga ccctgctcaa acccggagat accatcatat tcgaagccaa tggcaatctc 2340
atcgctccca tgtatgcttt cgctctctct cgcggattcg gttccggcat aatcactagc 2400
caagcctcca tgcacgagtg caataccaag tgtcaaactc cactgggagc cattcaaagc 2460
tccctgccct accaaaacat ccatcctgtg accattggag agtgccctaa atacgtgcgc 2520
agcgcgaaac tgcgcatggt gaccggtctc cgcaatatac cctccataca atcgcgcggc 2580
ctgtttggtg ctatcgctgg ctttatcgag ggaggatgga ctggaatgat cgacggctgg 2640
tatggctatc atcatcaaaa cgaacagggt tccggctacg ccgctgacca aaagtccact 2700
caaaacgcca ttaacggtat tacaaacaaa gtaaacaccg tgatagagaa aatgaatatc 2760
caattcactg ccgtgggcaa agagtttaac aagctggaga agcgcatgga aaatctgaac 2820
aaaaaagtcg atgatggctt cctcgacatc tggacttaca acgccgaact cctcgtgctg 2880
ctcgaaaacg agaggactct ggacttccac gactccaacg tgaagaacct gtacgagaaa 2940
gtcaaatccc aactcaagaa caacgccaaa gaaatcggca acggctgctt cgagttctac 3000
cacaaatgtg ataacgagtg tatggaatcc gtacggcaag gcacttacga ctaccccaaa 3060
tacagcgaag agagcaaatt gaaccgcgag aaagtggacg gcgtgaaact ggagtccatg 3120
ggcatctatc agcaccacca ccaccaccac taatag 3156

<210> 118
<211> 1557
<212> DNA
<213> Artificial Sequence

<220>
<223> ngHA0s

<400> 118
gacaccatct gcattggcta ccatgcccag cagtctacag atacagtcga taccgtcctg 60
gagaaacaag tcaccgtcac gcactccgtg aacctcctgg aggactccca taatggcaaa 120
ttgtgtcggc tgaaaggcat cgcccctctg caactcggca agtgcaatat tgcggggtgg 180
ctgttgggca acccagaatg tgaccccctc ctgcccgtcc gttcttggtc ttacatcgtg 240
gaaacaccta actccgagaa cggcatctgt tacccgggcg acttcattga ctacgaggag 300
ctccgcgaac agctgtcttc tgtctcttct ttcgaacgtt tcgaaatttt cccaaaggaa 360
tcctcctggc caaaccataa caccaacgga gtgaccgccg cctgtagcca tgagggcaag 420
tcttctttct accgtaatct gctgtggctg actgaaaaag agggttctta tcccaaactg 480
aagaactctt atgtcaacaa gaagggcaaa gaggtcctgg tgctgtgggg tatccaccac 540
ccccccaact ccaaggagca gcagaatctg taccaaaacg aaaatgctta cgtgtctgtc 600
gtgacatcta actacaaccg gcgcttcacg cccgaaatcg ccgagcgtcc caaagtgcgc 660
gatcaggccg gaaggatgaa ctactactgg accctgctca aacccggaga taccatcata 720
ttcgaagcca atggcaatct catcgctccc atgtatgctt tcgctctctc tcgcggattc 780
ggttccggca taatcactag ccaagcctcc atgcacgagt gcaataccaa gtgtcaaact 840
ccactgggag ccattcaaag ctccctgccc taccaaaaca tccatcctgt gaccattgga 900
gagtgcccta aatacgtgcg cagcgcgaaa ctgcgcatgg tgaccggtct ccgcaatata 960
ccctccatac aatcgcgcgg cctgtttggt gctatcgctg gctttatcga gggaggatgg 1020
actggaatga tcgacggctg gtatggctat catcatcaaa acgaacaggg ttccggctac 1080
gccgctgacc aaaagtccac tcaaaacgcc attaacggta ttacaaacaa agtaaacacc 1140
gtgatagaga aaatgaatat ccaattcact gccgtgggca aagagtttaa caagctggag 1200
aagcgcatgg aaaatctgaa caaaaaagtc gatgatggct tcctcgacat ctggacttac 1260
aacgccgaac tcctcgtgct gctcgaaaac gagaggactc tggacttcca cgactccaac 1320

gtgaagaacc tgtacgagaa agtcaaatcc caactcaaga acaacgccaa agaaatcggc 1380
aacggctgct tcgagttcta ccacaaatgt gataacgagt gtatggaatc cgtacggcaa 1440
ggcacttacg actaccccaa atacagcgaa gagagcaaat tgaaccgcga gaaagtggac 1500
ggcgtgaaac tggagtccat gggcatctat cagcaccacc accaccacca ctaatag 1557

<210> 119
<211> 2439
<212> DNA
<213> Artificial Sequence

<220>
<223> ngSTF2.HA1-1

<400> 119
atgaaattct tagtcaacgt tgcccttgtt tttatggtcg tatacatttc ttacatctat 60
gcggcccaag tgatcaacac caactccctg tccctgctca ctcaaaacaa cctccagaaa 120
tctcagtctg ctctcggtac tgctatcgag cgtctctcct ccggcctgag gatcaactcc 180
gctaaagatg acgccgccgg acaagctatc gctaaccgct ttactgccaa cattaaaggc 240
ctgacccaag cctctaggaa cgccaatgat ggcatttcta tcgctcaaac caccgaaggc 300
gccctgaacg aaatcaataa taacctgcaa cgtgtccgcg aactcgccgt ccagtccgcc 360
caatctacca actctcagtc tgacctggac tctatccaag ctgagatcac ccaaaggctc 420
aacgagatcg atcgtgtgtc tggtcaaact cagtttaacg gcgtcaaagt gctggctcaa 480
gacaacacac tcaccatcca agtaggagcc aatgacggcg agacaatcga tatcgacctg 540
aagcagatca attctcaaac tctgggcctg gactccctga acgtccaaaa ggcttacgac 600
gtgaaggaca ccgctgtgac aaccaaagcc tatgctaatc aaggaacaac cctggacgtg 660
tccggactcg atgacgccgc tatcaaggcc gctactggcg gcactcaggg aactgcctct 720
gtgaccggag gcgctgtgaa gttcgacgcc gataacaaca aatacttcgt cactattggt 780
ggtttcactg gcgctgacgc cgctaagaac ggtgactacg aggtcaacgt cgccactgac 840
ggaacagtga cactggccgc tggcgctacc aagaccacaa tgcctgctgg tgctactact 900
aagacagagg tgcaggaact caaggacacc cctgccgtgg tgtccgccga tgctaagaac 960
gctctcattg ctggtggtgt cgacgctacg gacgccaacg gagccgagct tgtgaaaatg 1020
tcctacaccg acaagaacgg aaagactatc gagggaggtt acgctctgaa ggctggcgat 1080
aagtactacg ccgctgatta tgacgaagct acaggcgcca ttaaagctaa aaccacatct 1140
tataccgctg ccgatggtac taccaagact gctgccaatc agctgggtgg agtcgatgga 1200
aaaaccgaag tggtcacaat cgacggaaag acctatcaag cctccaaggc tgctggccac 1260
gacttcaagg cccaacccga actggccgag gctgctgcta aaaccactga aaacccccctg 1320
cagaaaattg acgctgccct ggcccaagtg gatgccctcc gctccgacct cggcgctgtg 1380
caaaaccgct tcaactccgc tattacaaac ctcggcaata ccgtcaatca gctctctgaa 1440
gctaggtctc gtattgagga ttccgattac gctaccgagg tctcccagat gtcccgtgcc 1500
caaattctcc aacaagccgg cacctccgtc ctcgcccaag ccaatcaggt gccacagaat 1560
gtgctgagcc tgctggcttc tggatctgga tctggttctt cccataatgg caaattgtgt 1620
cggctgaaag gcatcgcccc tctgcaactc ggcaagtgca atattgcggg gtggctgttg 1680
ggcaacccag aatgtgaccc cctcctgccc gtccgttctt ggtcttacat cgtggaaaca 1740
cctaactccg agaacggcat ctgttacccg ggcgacttca ttgactacga ggagctccgc 1800
gaacagctgt cttctgtctc ttctttcgaa cgtttcgaaa ttttcccaaa ggaatcctcc 1860
tggccaaacc ataacaccaa cggagtgacc gccgcctgta gccatgaggg caagtcttct 1920
ttctaccgta atctgctgtg gctgactgaa aaagagggtt cttatcccaa actgaagaac 1980
tcttatgtca acaagaaggg caaagaggtc ctggtgctgt ggggtatcca ccacccccc 2040
aactccaagg agcagcagaa tctgtaccaa aacgaaaatg cttacgtgtc tgtcgtgaca 2100
tctaactaca accggcgctt cacgcccgaa atcgccgagc gtcccaaagt gcgcgatcag 2160
gccggaagga tgaactacta ctggaccctg ctcaaacccg gagataccat catattcgaa 2220
gccaatggca atctcatcgc tcccatgtat gctttcgctc tctctcgcgg attcggttcc 2280
ggcataatca ctagccaagc ctccatgcac gagtgcaata ccaagtgtca aactccactg 2340
ggagccattc aaagctccct gccctaccaa aacatccatc ctgtgaccat tggagagtgc 2400
cctaaatacg tgcgccacca ccaccaccac cactaatag 2439

<210> 120
<211> 840
<212> DNA

<213> Artificial Sequence

<220>
<223> ng HA1-1

<400> 120
tcccataatg gcaaattgtg tcggctgaaa ggcatcgccc ctctgcaact cggcaagtgc 60
aatattgcgg ggtggctgtt gggcaaccca gaatgtgacc ccctcctgcc cgtccgttct 120
tggtcttaca tcgtggaaac acctaactcc gagaacggca tctgttaccc gggcgacttc 180
attgactacg aggagctccg cgaacagctg tcttctgtct cttctttcga acgtttcgaa 240
attttcccaa aggaatcctc ctggccaaac cataacacca acggagtgac cgccgcctgt 300
agccatgagg gcaagtcttc tttctaccgt aatctgctgt ggctgactga aaaagagggt 360
tcttatccca aactgaagaa ctcttatgtc aacaagaagg gcaaagaggt cctggtgctg 420
tggggtatcc accacccccc caactccaag gagcagcaga atctgtacca aaacgaaaat 480
gcttacgtgt ctgtcgtgac atctaactac aaccggcgct tcacgcccga aatcgccgag 540
cgtcccaaag tgcgcgatca ggccggaagg atgaactact actggaccct gctcaaaccc 600
ggagatacca tcatattcga agccaatggc aatctcatcg ctcccatgta tgctttcgct 660
ctctctcgcg gattcggttc cggcataatc actagccaag cctccatgca cgagtgcaat 720
accaagtgtc aaactccact gggagccatt caaagctccc tgccctacca aaacatccat 780
cctgtgacca ttggagagtg ccctaaatac gtgcgccacc accaccacca ccactaatag 840

<210> 121
<211> 2292
<212> DNA
<213> Artificial Sequence

<220>
<223> ngSTF2.HA1-2

<400> 121
atgaaattct tagtcaacgt tgcccttgtt tttatggtcg tatacatttc ttacatctat 60
gcggcccaag tgatcaacac caactccctg tccctgctca ctcaaaacaa cctccagaaa 120
tctcagtctg ctctcggtac tgctatcgag cgtctctcct ccggcctgag gatcaactcc 180
gctaaagatg acgccgccgg acaagctatc gctaaccgct ttactgccaa cattaaaggc 240
ctgacccaag cctctaggaa cgccaatgat ggcatttcta tcgctcaaac caccgaaggc 300
gccctgaacg aaatcaataa taacctgcaa cgtgtccgcg aactcgccgt ccagtccgcc 360
caatctacca actctcagtc tgacctggac tctatccaag ctgagatcac ccaaaggctc 420
aacgagatcg atcgtgtgtc tggtcaaact cagtttaacg gcgtcaaagt gctggctcaa 480
gacaacacac tcaccatcca agtaggagcc aatgacggcg agacaatcga tatcgacctg 540
aagcagatca attctcaaac tctgggcctg gactccctga acgtccaaaa ggcttacgac 600
gtgaaggaca ccgctgtgac aaccaaagcc tatgctaatc aaggaacaac cctggacgtg 660
tccggactcg atgacgccgc tatcaaggcc gctactggcg gcactcaggg aactgcctct 720
gtgaccggag gcgctgtgaa gttcgacgcc gataacaaca aatacttcgt cactattggt 780
ggtttcactg gcgctgacgc cgctaagaac ggtgactacg aggtcaacgt cgccactgac 840
ggaacagtga cactggccgc tggcgctacc aagaccacaa tgcctgctgg tgctactact 900
aagacagagg tgcaggaact caaggacacc cctgccgtgg tgtccgccga tgctaagaac 960
gctctcattg ctggtggtgt cgacgctacg gacgccaacg gagccgagct tgtgaaaatg 1020
tcctacaccg acaagaacgg aaagactatc gagggaggtt acgctctgaa ggctggcgat 1080
aagtactacg ccgctgatta tgacgaagct acaggcgcca ttaaagctaa aaccacatct 1140
tataccgctg ccgatggtac taccaagact gctgccaatc agctgggtgg agtcgatgga 1200
aaaaccgaag tggtcacaat cgacggaaag acctatcaag cctccaaggc tgctggccac 1260
gacttcaagg cccaacccga actggccgag gctgctgcta aaaccactga aaaccccctg 1320
cagaaaattg acgctgccct ggcccaagtg gatgccctcc gctccgacct cggcgctgtg 1380
caaaaccgct tcaactccgc tattacaaac ctcggcaata ccgtcaatca gctctctgaa 1440
gctaggtctc gtattgagga ttccgattac gctaccgagg tctcccagat gtcccgtgcc 1500
caaattctcc aacaagccgg cacctccgtc ctcgcccaag ccaatcaggt gccacagaat 1560
gtgctgagcc tactagcatc aggttccggc tcaggttcca aaggcatagc gccactgcag 1620
ctcggcaaat gtaacatcgc aggttggctc cttggtaacc cggagtgcga ccccctcctc 1680

cctgtacgat cctggagtta tatcgtggag acccccaata gcgagaacgg aatttgctac 1740
ccaggagatt tcatagacta cgaggagttg cgcgagcagc tttcgtctgt gagcagcttc 1800
gagaggttcg aaatcttccc gaaggagagc agctggccga atcataacac taacggtgtg 1860
acagccgcct gcagtcatga aggaaagagt tcattctatc gcaacctgct gtggttgacg 1920
gagaaagagg gcagctaccc taagttgaag aactcctatg tgaacaaaaa aggcaaggag 1980
gttctggtgc tgtggggcat acaccacccc cccaatagca aggagcagca gaatctgtac 2040
caaaacgaga atgcctatgt gagcgtggtc actagtaact ataaccgtcg gttcactccc 2100
gagatcgccg agcgtccgaa ggtgagggac caggcaggcc ggatgaacta ctactggacc 2160
ctattgaagc caggggacac gattatcttc gaggcaaacg gaaacctcat agcgccgatg 2220
tacgcgttcg ccctgagccg tggctttgga tcggggatca tcacgtctca ccatcatcac 2280
caccattaat ag 2292

<210> 122
<211> 693
<212> DNA
<213> Artificial Sequence

<220>
<223> ng HA1-2

<400> 122
aaaggcatag cgccactgca gctcggcaaa tgtaacatcg caggttggct ccttggtaac 60
ccggagtgcg acccctcct ccctgtacga tcctggagtt atatcgtgga gacccccaat 120
agcgagaacg gaatttgcta cccaggagat ttcatagact acgaggagtt gcgcgagcag 180
ctttcgtctg tgagcagctt cgaaaggttc gagatattcc cgaaggagag cagctggccg 240
aatcataaca ctaacggtgt gacagccgcc tgcagtcatg aaggaaagag ttcattctat 300
cgcaacctgc tgtggttgac ggagaaagag ggcagctacc ctaagttgaa gaactcctat 360
gtgaacaaaa aaggcaagga ggttctggtg ctgtggggca tacaccaccc ccccaatagc 420
aaggagcagc agaatctgta ccaaaacgag aatgcctatg tgagcgtggt cactagtaac 480
tataaccgtc ggttcactcc cgagatcgcc gagcgtccga aggtgaggga ccaggcaggc 540
cggatgaact actactggac cctattgaag ccaggggaca cgattatctt cgaggcaaac 600
ggaaacctca tagcgccgat gtacgccttc gccctgagcc gtggctttgg atcggggatc 660
atcacgtctc accatcatca ccaccattaa tag 693

<210> 123
<211> 2292
<212> DNA
<213> Artificial Sequence

<220>
<223> ngSTF2.HA1-2mut

<400> 123
atgaaattct tagtcaacgt tgcccttgtt tttatggtcg tatacatttc ttacatctat 60
gcggcccaag tgatcaacac caactccctg tccctgctca ctcaaaacaa cctccagaaa 120
tctcagtctg ctctcggtac tgctatcgag cgtctctcct ccggcctgag gatcaactcc 180
gctaaagatg acgccgccgg acaagctatc gctaaccgct ttactgccaa cattaaaggc 240
ctgacccaag cctctaggaa cgccaatgat ggcatttcta tcgctcaaac caccgaaggc 300
gccctgaacg aaatcaataa taacctgcaa cgtgtccgcg aactcgccgt ccagtccgcc 360
caatctacca actctcagtc tgacctggac tctatccaag ctgagatcac ccaaaggctc 420
aacgagatcg atcgtgtgtc tggtcaaact cagtttaacg gcgtcaaagt gctggctcaa 480
gacaacacac tcaccatcca agtaggagcc aatgacggcg agacaatcga tatcgacctg 540
aagcagatca attctcaaac tctgggcctg gactccctga acgtccaaaa ggcttacgac 600
gtgaaggaca ccgctgtgac aaccaaagcc tatgctaatc aaggaacaac cctggacgtg 660
tccggactcg atgacgccgc tatcaaggcc gctactggcg gcactcaggg aactgcctct 720
gtgaccggag gcgctgtgaa gttcgacgcc gataacaaca aatacttcgt cactattggt 780
ggtttcactg gcgctgacgc cgctaagaac ggtgactacg aggtcaacgt cgccactgac 840
ggaacagtga cactggccgc tggcgctacc aagaccacaa tgcctgctgg tgctactact 900
aagacagagg tgcaggaact caaggacacc cctgccgtgg tgtccgccga tgctaagaac 960

gctctcattg ctggtggtgt cgacgctacg gacgccaacg gagccgagct tgtgaaaatg 1020
tcctacaccg acaagaacgg aaagactatc gagggaggtt acgctctgaa ggctggcgat 1080
aagtactacg ccgctgatta tgacgaagct acaggcgcca ttaaagctaa aaccacatct 1140
tataccgctg ccgatggtac taccaagact gctgccaatc agctgggtgg agtcgatgga 1200
aaaaccgaag tggtcacaat cgacggaaag acctatcaag cctccaaggc tgctggccac 1260
gacttcaagg cccaacccga actggccgag gctgctgcta aaaccactga aaaccccctg 1320
cagaaaattg acgctgccct ggcccaagtg gatgccctcc gctccgacct cggcgctgtg 1380
caaaaccgct tcaactccgc tattacaaac ctcggcaata ccgtcaatca gctctctgaa 1440
gctaggtctc gtattgagga ttccgattac gctaccgagg tctcccagat gtcccgtgcc 1500
caaattctcc aacaagccgg cacctccgtc ctcgcccaag ccaatcaggt gccacagaat 1560
gtgctgagcc tgctggcttc tggatctgga tctggttcta aaggcgctgc ccctctgcaa 1620
ctcggcaagt gcaatattgc ggggtggctg ttgggcaacc cagaatgtga ccccctcctg 1680
cccgtccgtt cttggtctga catcgctgaa acacctaact ccgagaacgg catctgttac 1740
ccgggcgact tcattgacta cgaggagctc cgcgaacagc tgtcttctgt ctcttctttc 1800
gaacgtttcg aaattttccc aaaggaatcc tcctggccaa accataacac caacggagtg 1860
accgccgcct gtagccatga gggcaagtct tctttctacc gtaatctgct gtggctgact 1920
gaaaaagagg gttcttatcc caaactgaag aactcttatg tcaacaagaa gggcaaagag 1980
gtcctggtgc tgtggggtat ccaccacccc cccaactcca aggagcagca gaatctgtac 2040
caaaacgaaa atgcttacgt gtctgtcgtg acatctaact acaaccggcg cttcacgccc 2100
gaaatcgccg agcgtcccaa agtgcgcgat caggccggaa ggatgaacta ctactggacc 2160
ctgctcaaac ccggagatac catcatattc gaagccaatg gcaatctcat cgctcccatg 2220
tatgctttcg ctctctctcg cggattcggt tccggcataa tcactagcca ccaccaccac 2280
caccactaat ag 2292

<210> 124
<211> 693
<212> DNA
<213> Artificial Sequence

<220>
<223> ngHA1-2mut

<400> 124
aaaggcgctg cccctctgca actcggcaag tgcaatattg cggggtggct gttgggcaac 60
ccagaatgtg acccctcct gcccgtccgt tcttggtctg acatcgctga aacacctaac 120
tccgagaacg gcatctgtta cccgggcgac ttcattgact acgaggagct ccgcgaacag 180
ctgtcttctg tctcttcttt cgaacgtttc gaaattttcc caaaggaatc ctcctggcca 240
aaccataaca ccaacggagt gaccgccgcc tgtagccatg agggcaagtc ttctttctac 300
cgtaatctgc tgtggctgac tgaaaaagag ggttcttatc ccaaactgaa gaactcttat 360
gtcaacaaga agggcaaaga ggtcctggtg ctgtggggta tccaccaccc ccccaactcc 420
aaggagcagc agaatctgta ccaaaacgaa aatgcttacg tgtctgtcgt gacatctaac 480
tacaaccggc gcttcacgcc cgaaatcgcc gagcgtccca aagtgcgcga tcaggccgga 540
aggatgaact actactggac cctgctcaaa cccggagata ccatcatatt cgaagccaat 600
ggcaatctca tcgctcccat gtatgctttc gctctctctc gcggattcgg ttccggcata 660
atcactagcc accaccacca ccaccactaa tag 693

<210> 125
<211> 2151
<212> DNA
<213> Artificial Sequence

<220>
<223> ngSTF2.HA1-3

<400> 125
atgaaattct tagtcaacgt tgcccttgtt tttatggtcg tatacatttc ttacatctat 60
gcggcccaag tgatcaacac caactccctg tccctgctca ctcaaaacaa cctccagaaa 120
tctcagtctg ctctcggtac tgctatcgag cgtctctcct ccggcctgag gatcaactcc 180
gctaaagatg acgccgccgg acaagctatc gctaaccgct ttactgccaa cattaaaggc 240

ctgacccaag cctctaggaa cgccaatgat ggcatttcta tcgctcaaac caccgaaggc 300
gccctgaacg aaatcaataa taacctgcaa cgtgtccgcg aactcgccgt ccagtccgcc 360
caatctacca actctcagtc tgacctggac tctatccaag ctgagatcac ccaaaggctc 420
aacgagatcg atcgtgtgtc tggtcaaact cagtttaacg gcgtcaaagt gctggctcaa 480
gacaacacac tcaccatcca agtaggagcc aatgacggcg agacaatcga tatcgacctg 540
aagcagatca attctcaaac tctgggcctg gactccctga acgtccaaaa ggcttacgac 600
gtgaaggaca ccgctgtgac aaccaaagcc tatgctaatc aaggaacaac cctggacgtg 660
tccggactcg atgacgccgc tatcaaggcc gctactggcg gcactcaggg aactgcctct 720
gtgaccggag gcgctgtgaa gttcgacgcc gataacaaca aatacttcgt cactattggt 780
ggtttcactg gcgctgacgc cgctaagaac ggtgactacg aggtcaacgt cgccactgac 840
ggaacagtga cactggccgc tggcgctacc aagaccacaa tgcctgctgg tgctactact 900
aagacagagg tgcaggaact caaggacacc cctgccgtgg tgtccgccga tgctaagaac 960
gctctcattg ctggtggtgt cgacgctacg gacgccaacg gagccgagct tgtgaaaatg 1020
tcctacaccg acaagaacgg aaagactatc gagggaggtt acgctctgaa ggctggcgat 1080
aagtactacg ccgctgatta tgacgaagct acaggcgcca ttaaagctaa aaccacatct 1140
tataccgctg ccgatggtac taccaagact gctgccaatc agctgggtgg agtcgatgga 1200
aaaaccgaag tggtcacaat cgacggaaag acctatcaag cctccaaggc tgctggccac 1260
gacttcaagg cccaacccga actggccgag gctgctgcta aaaccactga aaacccctg 1320
cagaaaattg acgctgccct ggcccaagtg gatgccctcc gctccgacct cggcgctgtg 1380
caaaaccgct tcaactccgc tattacaaac ctcggcaata ccgtcaatca gctctctgaa 1440
gctaggtctc gtattgagga ttccgattac gctaccgagg tctcccagat gtcccgtgcc 1500
caaattctcc aacaagccgg cacctccgtc ctcgcccaag ccaatcaggt gccacagaat 1560
gtgctgagcc tactagcatc aggttccggc tcaggttcca atagcgagaa cggaatttgc 1620
tacccaggag atttcataga ctacgaggag ttgcgcgagc agctttcgtc tgtgagcagc 1680
ttcgagaggt tcgaaatctt cccgaaggag agcagctggc cgaatcataa cactaacggt 1740
gtgacagccg cctgcagtca tgaaggaaag agttcattct atcgcaacct gctgtggttg 1800
acggagaagg agggcagcta ccctaagctg aagaactcct atgtgaacaa aaaaggcaag 1860
gaggttctgg tgctgtgggg catacaccac ccccccaata gcaaggagca gcagaatctg 1920
taccaaaacg agaatgccta tgtgagcgtg gtcactagta actataaccg tcggttcact 1980
cccgagatcg ccgagcgtcc gaaggtgagg gaccaggcag gccggatgaa ctactactgg 2040
accctattga agccagggga cacgattatc ttcgaggcaa acggaaacct catagcgccg 2100
atgtacgcct tcgccctgag ccgtggccac catcatcacc accattaata g 2151

<210> 126
<211> 552
<212> DNA
<213> Artificial Sequence

<220>
<223> ngHA1-3

<400> 126
aatagcgaga acggaatttg ctacccagga gatttcatag actacgagga gttgcgcgag 60
cagctttcgt ctgtgagcag cttcgaaagg ttcgagatat tcccgaagga gagcagctgg 120
ccgaatcata acactaacgg tgtgacagcc gcctgcagtc atgaaggaaa gagttcattc 180
tatcgcaacc tgctgtggtt gacggagaaa gagggcagct accctaagtt gaagaactcc 240
tatgtgaaca aaaaaggcaa ggaggttctg gtgctgtggg gcatacacca ccccccaat 300
agcaaggagc agcagaatct gtaccaaaac gagaatgcct atgtgagcgt ggtcactagt 360
aactataacc gtcggttcac tcccgagatc gccgagcgtc cgaaggtgag ggaccaggca 420
ggccggatga actactactg gaccctattg aagccagggg acacgattat cttcgaggca 480
aacggaaacc tcatagcgcc gatgtacgcc ttcgccctga gccgtggcca ccatcatcac 540
caccattaat ag 552

<210> 127
<211> 2151
<212> DNA
<213> Artificial Sequence

<220>

<223> ngSTF2.HA1-3mut

<400> 127
atgaaattct tagtcaacgt tgcccttgtt tttatggtcg tatacatttc ttacatctat 60
gcggcccaag tgatcaacac caactccctg tccctgctca ctcaaaacaa cctccagaaa 120
tctcagtctg ctctcggtac tgctatcgag cgtctctcct ccggcctgag gatcaactcc 180
gctaaagatg acgccgccgg acaagctatc gctaaccgct ttactgccaa cattaaaggc 240
ctgacccaag cctctaggaa cgccaatgat ggcatttcta tcgctcaaac caccgaaggc 300
gccctgaacg aaatcaataa taacctgcaa cgtgtccgcg aactcgccgt ccagtccgcc 360
caatctacca actctcagtc tgacctggac tctatccaag ctgagatcac ccaaaggctc 420
aacgagatcg atcgtgtgtc tggtcaaact cagtttaacg gcgtcaaagt gctggctcaa 480
gacaacacac tcaccatcca agtaggagcc aatgacggcg agacaatcga tatcgacctg 540
aagcagatca attctcaaac tctgggcctg gactccctga acgtccaaaa ggcttacgac 600
gtgaaggaca ccgctgtgac aaccaaagcc tatgctaatc aaggaacaac cctggacgtg 660
tccggactcg atgacgccgc tatcaaggcc gctactggcg gcactcaggg aactgcctct 720
gtgaccggag gcgctgtgaa gttcgacgcc gataacaaca aatacttcgt cactattggt 780
ggtttcactg gcgctgacgc cgctaagaac ggtgactacg aggtcaacgt cgccactgac 840
ggaacagtga cactggccgc tggcgctacc aagaccacaa tgcctgctgg tgctactact 900
aagacagagg tgcaggaact caaggacacc cctgccgtgg tgtccgccga tgctaagaac 960
gctctcattg ctggtggtgt cgacgctacg gacgccaacg gagccgagct tgtgaaaatg 1020
tcctacaccg acaagaacgg aaagactatc gagggaggtt acgctctgaa ggctggcgat 1080
aagtactacg ccgctgatta tgacgaagct acaggcgcca ttaaagctaa aaccacatct 1140
tataccgctg ccgatggtac taccaagact gctgccaatc agctgggtgg agtcgatgga 1200
aaaaccgaag tggtcacaat cgacggaaag acctatcaag cctccaaggc tgctggccac 1260
gacttcaagg cccaacccga actggccgag gctgctgcta aaaccactga aaacccccctg 1320
cagaaaattg acgctgccct ggcccaagtg gatgccctcc gctccgacct cggcgctgtg 1380
caaaaccgct tcaactccgc tattacaaac ctcggcaata ccgtcaatca gctctctgaa 1440
gctaggtctc gtattgagga ttccgattac gctaccgagg tctcccagat gtcccgtgcc 1500
caaattctcc aacaagccgg cacctccgtc ctcgcccaag ccaatcaggt gccacagaat 1560
gtgctgagcc tgctggcttc tggatctgga tctggttcta actccgagaa cgaaatctgt 1620
tacccgggcg acttcattga caaagaggag ctccgcgaac agctgtcttc tgtctcttct 1680
ttcgaacgtt tcgaaatttt cccaaaggaa tcctcctggc caaaccataa caccaacgga 1740
gtgaccgccg cctgtagcca tgagggcaag tcttctttct accgtaatct gctgtggctg 1800
actgaaaaag agggttctta tcccaaactg aagaactctt atgtcaacaa gaagggcaaa 1860
gaggtcctgg tgctgtgggg tatccaccac ccccccaact ccaaggagca gcagaatctg 1920
taccaaaacg aaaatgctta cgtgtctgtc gtgacatcta actacaaccg gcgcttcacg 1980
cccgaaatcg ccgagcgtcc caaagtgcgc gatcaggccg gaaggatgaa ctactactgg 2040
accctgctca aacccggaga taccatcata ttcgaagcca atggcaatct catcgctccc 2100
atgtatgctg ctgctctctc tcgcggacac caccaccacc accactaata g 2151

<210> 128
<211> 552
<212> DNA
<213> Artificial Sequence

<220>
<223> ngHA1-3mut

<400> 128
aactccgaga acgaaatctg ttacccgggc gacttcattg acaaagagga gctccgcgaa 60
cagctgtctt ctgtctcttc tttcgaacgt ttcgaaattt tcccaaagga atcctcctgg 120
ccaaaccata acaccaacgg agtgaccgcc gcctgtagcc atgagggcaa gtcttctttc 180
taccgtaatc tgctgtggct gactgaaaaa gagggttctt atcccaaact gaagaactct 240
tatgtcaaca agaagggcaa agaggtcctg gtgctgtggg gtatccacca ccccccaac 300
tccaaggagc agcagaatct gtaccaaaac gaaaatgctt acgtgtctgt cgtgacatct 360
aactacaacc ggcgcttcac gcccgaaatc gccgagcgtc ccaaagtgcg cgatcaggcc 420
ggaaggatga actactactg gaccctgctc aaacccggag ataccatcat attcgaagcc 480
aatggcaatc tcatcgctcc catgtatgct gctgctctct ctcgcggaca ccaccaccac 540
caccactaat ag 552

<210> 129
<211> 2439
<212> DNA
<213> Artificial Sequence

<220>
<223> wtSTF2.HA1-1ng

<400> 129
atgaaattct tagtcaacgt tgcccttgtt tttatggtcg tatacatttc ttacatctat 60
gcggcccagg ttatcaatac caactccctg tcgttgctca cccaaaataa ccttaataaa 120
agccagagcg cactgggaac cgccatagaa cgcctctcaa gcggcctccg gatcaattct 180
gcaaaagacg acgccgccgg tcaggccatc gcaaaccgct ttaccgccaa tatcaaggga 240
ctgacgcagg cttcgaggaa tgctaacgat ggaataagca tcgctcaaac cacggagggc 300
gccctgaacg agatcaacaa caacctacag cgcgtcaggg agctcgcagt gcagtccgcc 360
aattcgacca actcgcagtc ggacctggac tcgatccaag ccgaaatcac ccagcgcctg 420
aatgagattg accgggtgag cggtcagaca cagtttaacg gcgtgaaggt acttgcacag 480
gataacacac ttacgataca ggtgggcgcc aacgatggtg aaaccataga cattgatctc 540
aaacagatta acagccagac gctcgggttg gatagcctga atgtgcaaaa ggcgtacgac 600
gtgaaagaca cggcggtcac taccaaagcc tacgctaaca atggcactac cttggatgtg 660
agcggattgg atgatgcagc aatcaaggct gctaccggcg gtacgaacgg aaccgcgtcc 720
gtgaccggcg gtgccgtgaa gttcgatgct gacaacaata agtatttcgt caccattgga 780
ggctttactg gcgccgacgc agcaaagaac ggcgactatg aagtgaacgt ggcaaccgat 840
ggaaccgtga cgctggccgc tggtgccacc aagaccacca tgccagccgg cgccacaact 900
aagaccgagg tgcaggagtt aaaggacacc cccgcggtgg ttagcgcaga tgccaaaaac 960
gccttgatcg ccggcggagt ggatgcaact gatgctaatg gtgcggagct ggttaaaatg 1020
tcgtatacag acaagaatgg taagacgatc gagggcggtt atgcccttaa ggcaggagat 1080
aagtattacg ctgctgatta cgatgaggcg acgggagcta ttaaggccaa gacaacgtca 1140
tacacggcgg cggacggaac gactaagacg gctgccaatc agttgggagg ggttgacggg 1200
aagacagagg tcgttacgat cgatggcaag acatacaacg cttccaaggc cgctggccac 1260
gatttcaaag ctcaacccga actggccgag gccgcggcga aaacaactga gaacccgttg 1320
cagaagattg atgcggccct ggcgcaagta gatgccctgc gctcagacct gggcgccgtt 1380
caaaatcgat tcaattccgc gattacaaac ctgggcaata cagtaaacaa tctatccgag 1440
gccagatccc gcattgaaga ctccgactac gcgacagaag taagtaacat gagtcgtgcc 1500
cagattctgc agcaggccgg cactagtgtc ctggcccagg ccaatcaagt cccgcagaat 1560
gtgctgagcc tactagcatc tggatctgga tctggttctt cccataatgg caaattgtgt 1620
cggctgaaag gcatcgcccc tctgcaactc ggcaagtgca atattgcggg gtggctgttg 1680
ggcaacccag aatgtgaccc cctcctgccc gtccgttctt ggtcttacat cgtggaaaca 1740
cctaactccg agaacggcat ctgttacccg ggcgacttca ttgactacga ggagctccgc 1800
gaacagctgt cttctgtctc ttctttcgaa cgtttcgaaa ttttcccaaa ggaatcctcc 1860
tggccaaacc ataacaccaa cggagtgacc gccgcctgta gccatgaggg caagtcttct 1920
ttctaccgta atctgctgtg gctgactgaa aaagagggtt cttatcccaa actgaagaac 1980
tcttatgtca acaagaaggg caaagaggtc ctggtgctgt ggggtatcca ccacccccc 2040
aactccaagg agcagcagaa tctgtaccaa aacgaaaatg cttacgtgtc tgtcgtgaca 2100
tctaactaca accggcgctt cacgcccgaa atcgccgagc gtcccaaagt gcgcgatcag 2160
gccggaagga tgaactacta ctggaccctg ctcaaacccg gagataccat catattcgaa 2220
gccaatggca atctcatcgc tcccatgtat gctttcgctc tctctcgcgg attcggttcc 2280
ggcataatca ctagccaagc ctccatgcac gagtgcaata ccaagtgtca aactccactg 2340
ggagccattc aaagctccct gccctaccaa aacatccatc ctgtgaccat tggagagtgc 2400
cctaaatacg tgcgccacca ccaccaccac cactaatag 2439

<210> 130
<211> 840
<212> DNA
<213> Artificial Sequence

<220>
<223> ngHA1-1

<400> 130
tcccataatg gcaaattgtg tcggctgaaa ggcatcgccc ctctgcaact cggcaagtgc 60
aatattgcgg ggtggctgtt gggcaaccca gaatgtgacc ccctcctgcc cgtccgttct 120
tggtcttaca tcgtggaaac acctaactcc gagaacggca tctgttaccc gggcgacttc 180
attgactacg aggagctccg cgaacagctg tcttctgtct cttctttcga acgtttcgaa 240
attttcccaa aggaatcctc ctggccaaac cataacacca acggagtgac cgccgcctgt 300
agccatgagg gcaagtcttc tttctaccgt aatctgctgt ggctgactga aaaagagggt 360
tcttatccca aactgaagaa ctcttatgtc aacaagaagg gcaaagaggt cctggtgctg 420
tggggtatcc accacccccc caactccaag gagcagcaga atctgtacca aaacgaaaat 480
gcttacgtgt ctgtcgtgac atctaactac aaccggcgct tcacgcccga aatcgccgag 540
cgtcccaaag tgcgcgatca ggccggaagg atgaactact actggaccct gctcaaaccc 600
ggagatacca tcatattcga agccaatggc aatctcatcg ctcccatgta tgctttcgct 660
ctctctcgcg gattcggttc cggcataatc actagccaag cctccatgca cgagtgcaat 720
accaagtgtc aaactccact gggagccatt caaagctccc tgccctacca aaacatccat 780
cctgtgacca ttggagagtg ccctaaatac gtgcgccacc accaccacca ccactaatag 840

<210> 131
<211> 2439
<212> DNA
<213> Artificial Sequence

<220>
<223> ngSTF2.HA1-1wt

<400> 131
atgaaattct tagtcaacgt tgcccttgtt tttatggtcg tatacatttc ttacatctat 60
gcggcccaag tgatcaacac caactccctg tccctgctca ctcaaaacaa cctccagaaa 120
tctcagtctg ctctcggtac tgctatcgag cgtctctcct ccggcctgag gatcaactcc 180
gctaaagatg acgccgccgg acaagctatc gctaaccgct ttactgccaa cattaaaggc 240
ctgacccaag cctctaggaa cgccaatgat ggcatttcta tcgctcaaac caccgaaggc 300
gccctgaacg aaatcaataa taacctgcaa cgtgtccgcg aactcgccgt ccagtccgcc 360
caatctacca actctcagtc tgacctggac tctatccaag ctgagatcac ccaaaggctc 420
aacgagatcg atcgtgtgtc tggtcaaact cagtttaacg gcgtcaaagt gctggctcaa 480
gacaacacac tcaccatcca agtaggagcc aatgacggcg agacaatcga tatcgacctg 540
aagcagatca attctcaaac tctgggcctg gactccctga acgtccaaaa ggcttacgac 600
gtgaaggaca ccgctgtgac aaccaaagcc tatgctaatc aaggaacaac cctggacgtg 660
tccggactcg atgacgccgc tatcaaggcc gctactggcg gcactcaggg aactgcctct 720
gtgaccggag gcgctgtgaa gttcgacgcc gataacaaca aatacttcgt cactattggt 780
ggtttcactg gcgctgacgc cgctaagaac ggtgactacg aggtcaacgt cgccactgac 840
ggaacagtga cactggccgc tggcgctacc aagaccacaa tgcctgctgg tgctactact 900
aagacagagg tgcaggaact caaggacacc cctgccgtgg tgtccgccga tgctaagaac 960
gctctcattg ctggtggtgt cgacgctacg gacgccaacg gagccgagct tgtgaaaatg 1020
tcctacaccg acaagaacgg aaagactatc gagggaggtt acgctctgaa ggctggcgat 1080
aagtactacg ccgctgatta tgacgaagct acaggcgcca ttaaagctaa aaccacatct 1140
tataccgctg ccgatggtac taccaagact gctgccaatc agctgggtgg agtcgatgga 1200
aaaaccgaag tggtcacaat cgacggaaag acctatcaag cctccaaggc tgctggccac 1260
gacttcaagg cccaacccga actggccgag gctgctgcta aaaccactga aaaccccctg 1320
cagaaaattg acgctgccct ggcccaagtg gatgccctcc gctccgacct cggcgctgtg 1380
caaaaccgct tcaactccgc tattacaaac ctcggcaata ccgtcaatca gctctctgaa 1440
gctaggtctc gtattgagga ttccgattac gctaccgagg tctcccagat gtcccgtgcc 1500
caaattctcc aacaagccgg cacctccgtc ctcgcccaag ccaatcaggt gccacagaat 1560
gtgctgagcc tactagcatc aggttccggc tcaggttcca gccacaacgg aaagctgtgc 1620
cgtctgaaag gcatagcgcc actgcagctc ggcaaatgta acatcgcagg ttggctcctt 1680
ggtaacccgg agtgcgaccc cctcctccct gtacgatcct ggagttatat cgtggagacc 1740
cccaatagcg agaacggaat ttgctaccca ggagatttca tagactacga ggagttgcgc 1800
gagcagcttt cgtctgtgag cagcttcgag aggttcgaaa tcttcccgaa ggagagcagc 1860
tggccgaatc ataacactaa cggtgtgaca gccgcctgca gtcatgaagg aaagagttca 1920

ttctatcgca acctgctgtg gttgacggag aaagagggca gctaccctaa gttgaagaac 1980
tcctatgtga acaaaaaagg caaggaggtt ctggtgctgt ggggcataca ccacccccc 2040
aatagcaagg agcagcagaa tctgtaccaa aacgagaatg cctatgtgag cgtggtcact 2100
agtaactata accgtcggtt cactcccgag atcgccgagc gtccgaaggt gagggaccag 2160
gcaggccgga tgaactacta ctggacccta ttgaagccag gggacacgat tatcttcgag 2220
gcaaacggaa acctcatagc gccgatgtac gcgttcgccc tgagccgcgg ctttggatcg 2280
gggatcatca cgtctaacgc ctcgatgcac gaatgtaata ccaaatgcca gaccccactg 2340
ggtgctatca actcgtcctt accctatcaa aatatacatc cggtcaccat aggcgagtgt 2400
cccaaatatg tcagacacca tcatcaccac cattaatag 2439

<210> 132
<211> 840
<212> DNA
<213> Artificial Sequence

<220>
<223> HA1-1 PR8

<400> 132
agccacaacg gaaagctgtg ccgtctgaaa ggcatagcgc cactgcagct cggcaaatgt 60
aacatcgcag gttggctcct tggtaacccg gagtgcgacc ccctcctccc tgtacgatcc 120
tggagttata tcgtggagac ccccaatagc gagaacggaa tttgctaccc aggagatttc 180
atagactacg aggagttgcg cgagcagctt tcgtctgtga gcagcttcga aaggttcgag 240
atattcccga aggagagcag ctggccgaat cataacacta acggtgtgac agccgcctgc 300
agtcatgaag gaaagagttc attctatcgc aacctgctgt ggttgacgga gaaagagggc 360
agctacccta agttgaagaa ctcctatgtg aacaaaaaag gcaaggaggt tctggtgctg 420
tggggcatac accacccccc caatagcaag gagcagcaga atctgtacca aaacgagaat 480
gcctatgtga gcgtggtcac tagtaactat aaccgtcggt tcactcccga gatcgccgag 540
cgtccgaagg tgagggacca ggcaggccgg atgaactact actggaccct attgaagcca 600
ggggacacga ttatcttcga ggcaaacgga aacctcatag cgccgatgta cgccttcgcc 660
ctgagccgtg gctttggatc ggggatcatc acgtctaacg cctcgatgca cgaatgtaat 720
accaaatgcc agaccccact gggtgctatc aactcgtcct taccctatca aaatatacat 780
ccggtcacca taggcgagtg tcccaaatat gtcagacacc atcatcacca ccattaatag 840

<210> 133
<211> 903
<212> DNA
<213> Artificial Sequence

<220>
<223> HA1-1 PR8

<400> 133
atgaaattct tagtcaacgt tgcccttgtt tttatggtcg tatacatttc ttacatctat 60
gcgagccaca acggaaagct gtgccgtctg aaaggcatag cgccactgca gctcggcaaa 120
tgtaacatcg caggttggct ccttggtaac ccggagtgcg acccctcct ccctgtacga 180
tcctggagtt atatcgtgga gacccccaat agcgagaacg gaatttgcta cccaggagat 240
ttcatagact acgaggagtt gcgcgagcag ctttcgtctg tgagcagctt cgaaaggttc 300
gagatattcc cgaaggagag cagctggccg aatcataaca ctaacggtgt gacagccgcc 360
tgcagtcatg aaggaaagag ttcattctat cgcaacctgc tgtggttgac ggagaaagag 420
ggcagctacc ctaagttgaa gaactcctat gtgaacaaaa aaggcaagga ggttctggtg 480
ctgtggggca tacaccaccc ccccaatagc aaggagcagc agaatctgta ccaaaacgag 540
aatgcctatg tgagcgtggt cactagtaac tataaccgtc ggttcactcc cgagatcgcc 600
gagcgtccga aggtgaggga ccaggcaggc cggatgaact actactggac cctattgaag 660
ccaggggaca cgattatctt cgaggcaaac ggaaacctca tagcgccgat gtacgccttc 720
gccctgagcc gtggctttgg atcggggatc atcacgtcta acgcctcgat gcacgaatgt 780
aataccaaat gccagacccc actgggtgct atcaactcgt ccttacccta tcaaaatata 840
catccggtca ccataggcga gtgtcccaaa tatgtcagac accatcatca ccaccattaa 900

tag 903

<210> 134
<211> 105
<212> DNA
<213> Artificial Sequence

<220>
<223> FOR PRIMER

<400> 134
aggcagatct atgaaattct tagtcaacgt tgcccttgtt tttatggtcg tatacatttc 60
ttacatctat gcgagccaca acggaaagct gtgccgtctg aaagg 105

<210> 135
<211> 55
<212> DNA
<213> Artificial Sequence

<220>
<223> REV PRIMER

<400> 135
acctgcatgc ctattaatgg tggtgatgat ggtgtctgac atatttggga cactc 55

<210> 136
<211> 1620
<212> DNA
<213> Artificial Sequence

<220>
<223> HA0s PR8

<400> 136
atgaaattct tagtcaacgt tgcccttgtt tttatggtcg tatacatttc ttacatctat 60
gcggacacca tttgcattgg ataccatgca aacaactcaa ccgatactgt tgataccgtc 120
cttgagaaga acgttaccgt cacgcactcg gtcaacctat tagaggatag ccacaacgga 180
aagctgtgcc gtctgaaagg catagcgcca ctgcagctcg gcaaatgtaa catcgcaggt 240
tggctccttg gtaacccgga gtgcgacccc ctcctccctg tacgatcctg gagttatatc 300
gtggagaccc ccaatagcga gaacggaatt tgctacccag gagatttcat agactacgag 360
gagttgcgcg agcagctttc gtctgtgagc agcttcgaaa ggttcgagat attcccgaag 420
gagagcagct ggccgaatca taacactaac ggtgtgacag ccgcctgcag tcatgaagga 480
aagagttcat tctatcgcaa cctgctgtgg ttgacggaga aagagggcag ctaccctaag 540
ttgaagaact cctatgtgaa caaaaaaggc aaggaggttc tggtgctgtg gggcatacac 600
cacccccca atagcaagga gcagcagaat ctgtaccaaa acgagaatgc ctatgtgagc 660
gtggtcacta gtaactataa ccgtcggttc actcccgaga tcgccgagcg tccgaaggtg 720
agggaccagg caggccggat gaactactac tggaccctat tgaagccagg ggacacgatt 780
atcttcgagg caaacggaaa cctcatagcg ccgatgtacg ccttcgccct gagccgtggc 840
tttggatcgg ggatcatcac gtctaacgcc tcgatgcacg aatgtaatac caaatgccag 900
accccactgg gtgctatcaa ctcgtcctta ccctatcaaa atatacatcc ggtcaccata 960
ggcgagtgtc ccaaatatgt cagatccgcc aagttgcgga tggtgaccgg cctccgcaat 1020
attcctagta ttcagtcacg cggcttgttc ggcgccatcg ctggtttcat cgaaggcggg 1080
tggacaggca tgattgatgg ctggtatggc tatcaccacc agaacgagca gggctcgggc 1140
tacgcggctg accagaagtc gactcagaat gccatcaatg gcatcacgaa taaggtgaac 1200
acggtcattg aaaagatgaa cattcaattt acagccgtag gaaaagagtt taataaactg 1260
gaaaaaagaa tggagaatct gaataagaag gtggacgacg gatttttgga catctggacg 1320
tacaacgccg agctgctggt tctgctggaa aatgagcgaa cactggattt tcatgattct 1380
aacgtaaaga atctgtacga gaaggtgaag tcccaactaa agaataatgc caaggaaatc 1440
ggaaatggat gctttgagtt ttaccacaag tgcgataatg agtgcatgga atccgtgcga 1500

aatggtacat acgattaccc aaagtactcc gaagaatcca agctaaatcg cgaaaaggtt 1560
gatggtgtta aacttgaatc catgggtatt taccaacacc atcatcacca ccattaatag 1620

<210> 137
<211> 885
<212> DNA
<213> Artificial Sequence

<220>
<223> HA1-1 PR8 (no his tag)

<400> 137
atgaaattct tagtcaacgt tgcccttgtt tttatggtcg tatacatttc ttacatctat 60
gcgagccaca acggaaagct gtgccgtctg aaaggcatag cgccactgca gctcggcaaa 120
tgtaacatcg caggttggct ccttggtaac ccggagtgcg acccccctcct ccctgtacga 180
tcctggagtt atatcgtgga gacccccaat agcgagaacg gaatttgcta cccaggagat 240
ttcatagact acgaggagtt gcgcgagcag ctttcgtctg tgagcagctt cgaaaggttc 300
gagatattcc cgaaggagag cagctggccg aatcataaca ctaacggtgt gacagccgcc 360
tgcagtcatg aaggaaagag ttcattctat cgcaacctgc tgtggttgac ggagaaagag 420
ggcagctacc ctaagttgaa gaactcctat gtgaacaaaa aaggcaagga ggttctggtg 480
ctgtggggca tacaccaccc ccccaatagc aaggagcagc agaatctgta ccaaaacgag 540
aatgcctatg tgagcgtggt cactagtaac tataaccgtc ggttcactcc cgagatcgcc 600
gagcgtccga aggtgaggga ccaggcaggc cggatgaact actactggac cctattgaag 660
ccaggggaca cgattatctt cgaggcaaac ggaaacctca tagcgccgat gtacgccttc 720
gccctgagcc gtggctttgg atcggggatc atcacgtcta acgcctcgat gcacgaatgt 780
aataccaaat gccagacccc actgggtgct atcaactcgt ccttacccta tcaaaatata 840
catccggtca ccataggcga gtgtcccaaa tatgtcagat aatag 885

<210> 138
<211> 45
<212> DNA
<213> Artificial Sequence

<220>
<223> REV PRIMER

<400> 138
acctgcatgc ctattatctg acatatttgg gacactcgcc tatgg 45

<210> 139
<211> 903
<212> DNA
<213> Artificial Sequence

<220>
<223> HA1-1 NC

<400> 139
atgaaattct tagtcaacgt tgcccttgtt tttatggtcg tatacatttc ttacatctat 60
gcgagccaca atggtaagct gtgtctgttg aagggaattg cacccctcca actgggcaac 120
tgttcggtcg ctggttggat cctcggaaac ccagaatgcg agctccttat cagtaaggaa 180
tcttggtctt atattgtcga aaccccgaac cccgagaacg gaacatgcta cccgggttac 240
tttgctgatt acgaagagct tcgcgagcaa ctcagctccg tatcctcctt cgagcgcttc 300
gagatttttc ccaaagagtc cagctggcca aatcataccg tcaccggcgt gtcggcctcc 360
tgttcccaca acggaaagtc tagcttctat agaaatcttc tctggctgac gggtaagaat 420
ggtctttacc ccaatttgag caagtcctac gtcaacaaca aagaaaagga agttctggta 480
ttgtggggtg tgcaccaccc tccgaacatc ggcaatcagc gcgccctgta tcacacagag 540
aacgcgtatg tttccgttgt ctcctcacat tactcgaggc gcttcactcc tgaaatagct 600

aagcgtccga aagtgcgtga ccaggaggga cgtatcaact attattggac gctgttggag 660
ccaggcgata caattatctt cgaggctaac ggtaacctta tcgctccctg gtacgccttc 720
gccctgtcgc gtggtttcgg tagtggaata atcactagta atgctcctat ggacgagtgt 780
gacgctaagt gccaaacacc tcagggcgct atcaatagct cccttccatt ccagaacgtc 840
catccggtta ccattggaga gtgtccaaag tacgtgagac accatcatca ccatcactaa 900
tag 903

<210> 140
<211> 103
<212> DNA
<213> Artificial Sequence

<220>
<223> FOR PRIMER

<400> 140
aggcagatct atgaaattct tagtcaacgt tgcccttgtt tttatggtcg tatacatttc 60
ttacatctat gcgagccaca atggtaagct gtgtctgttg aag 103

<210> 141
<211> 63
<212> DNA
<213> Artificial Sequence

<220>
<223> REV PRIMER

<400> 141
agtcgcatgc ctattaatgg tggtgatgat ggtgtctcac gtactttgga cactctccaa 60
tgg 63

<210> 142
<211> 1632
<212> DNA
<213> Artificial Sequence

<220>
<223> HA0s NC

<400> 142
agatctatga agtttctcgt gaacgttgca ctggtgttta tggtagtcta tatatcgtac 60
atttatgctg acacgatttg catcggttac catgcgaaca actccacgga taccgtggac 120
acagtgttgg aaaagaacgt gaccgtcacg cactccgtaa accttctgga ggacagccac 180
aatggtaagc tgtgtctgtt gaagggaatt gcacccctcc aactgggcaa ctgttcggtc 240
gctggttgga tcctcggaaa cccagaatgc gagctcctta tcagtaagga atcttggtct 300
tatattgtcg aaaccccgaa ccccgagaac ggaacatgct acccgggtta ctttgctgat 360
tacgaagagc ttcgcgagca actcagctcc gtatcctcct tcgagcgctt cgagattttt 420
cccaaagagt ccagctggcc aaatcatacc gtcaccggcg tgtcggcctc ctgttcccac 480
aacggaaagt ctagcttcta tagaaatctt ctctggctga cgggtaagaa tggtctttac 540
cccaatttga gcaagtccta cgtcaacaac aaagaaaagg aagttctggt attgtggggt 600
gtgcaccacc ctccgaacat cggcaatcag cgcgccctgt atcacacaga gaacgcgtat 660
gtttccgttg tctcctcaca ttactcgagg cgcttcactc ctgaaatagc taagcgtccg 720
aaagtgcgtg accaggaggg acgtatcaac tattattgga cgctgttgga gccaggcgat 780
acaattatct tcgaggctaa cggtaacctt atcgctccct ggtacgcctt cgccctgtcg 840
cgtggtttcg gtagtggaat aatcactagt aatgctccta tggacgagtg tgacgctaag 900
tgccaaacac ctcagggcgc tatcaatagc tcccttccat tccagaacgt ccatccggtt 960
accattggag agtgtccaaa gtacgtgaga tcggccaaac ttcgcatggt cacgggtctg 1020
cgcaacatcc cgtcaatcca atctaggggc ctcttcggcg ctatcgccgg tttcattgag 1080
ggcggttgga ctggaatggt tgacggatgg tacggctatc atcaccagaa cgaacaaggt 1140

tccggttacg ctgctgacca gaaatctact cagaacgcga tcaatggtat cacgaacaag 1200
gtgaacagcg tcattgaaaa gatgaatact cagtttacag ccgtgggcaa agagttcaat 1260
aaactcgaga gacgtatgga aaacctcaat aagaaggtgg atgacggctt cctggacatt 1320
tggacttaca acgccgagct gctggtcctg ctcgagaacg agagaaccct tgacttccac 1380
gacagcaacg tcaagaacct gtacgagaag gtgaaaagtc aacttaaaaa caatgccaag 1440
gagattggta acggctgctt cgaattctac cacaagtgta ataatgagtg catggaatcc 1500
gttaagaacg gcacctacga ttaccctaaa tactcagagg agtccaagct taaccgcgag 1560
aagatcgacg gcgtaaaact ggaaagcatg ggcgtatacc aacaccatca tcaccatcac 1620
taataggcat gc 1632

<210> 143
<211> 909
<212> DNA
<213> Artificial Sequence

<220>
<223> HA1-1 VN

<400> 143
atgaaattct tagtcaacgt tgcccttgtt tttatggtcg tatacatttc ttacatctat 60
gcggagaaga aacacaacgg taagctttgc gacttggatg gagtcaagcc cctcatactt 120
agagattgta gtgtagccgg ttggctgctc ggtaacccaa tgtgcgatga gttcatcaat 180
gttcccgaat ggtcatatat cgtcgaaaaa gctaatcctg tcaacgacct gtgctacccc 240
ggtgatttca atgactatga agaactgaag cacctgctct cccgcatcaa ccatttcgag 300
aaaatccaga tcattcccaa gagttcctgg tctagccatg aggctagtct gggtgtctca 360
tccgcctgcc catatcaggg taaaagttct ttctttagga acgtagtatg gttgataaag 420
aaaaactcta catacccgac catcaagcgc tcttacaaca atacgaacca agaggatctg 480
cttgtccttt ggggaatcca tcatcctaat gatgctgccg aacagactaa gctctaccaa 540
aaccctacca cttatatttc cgtcggcacc tctactctga accagcgcct tgtgcccagg 600
atcgctacga gatcaaaagt caacggccaa tcgggccgca tggaattctt ctggacgatc 660
ctgaagccta atgacgctat caacttcgag tcaaatggaa actttatcgc tcccgagtac 720
gcttacaaga tcgtcaagaa gggcgactcc acgattatga agtcagagtt ggagtacggc 780
aactgcaaca caaagtgcca aactcctatg ggcgctataa attcttcaat gccgttccac 840
aacatccatc cgctcacgat cggtgagtgc ccgaaatatg taaagcacca tcaccaccat 900
cactaatag 909

<210> 144
<211> 99
<212> DNA
<213> Artificial Sequence

<220>
<223> FOR PRIMER

<400> 144
aggcagatct atgaaattct tagtcaacgt tgcccttgtt tttatggtcg tatacatttc 60
ttacatctat gcggagaaga aacacaacgg taagctttg 99

<210> 145
<211> 58
<212> DNA
<213> Artificial Sequence

<220>
<223> REV PRIMER

<400> 145
agtcgcatgc ctattatggt ggtgatgatg gtgctttaca tatttcgggc actcaccg 58

<210> 146
<211> 1644
<212> DNA
<213> Artificial Sequence

<220>
<223> HA0s VN

<400> 146
agatctatga aattcttggt taatgtagcc ctggtgttta tggtagtgta catttcatac 60
atttatgctg atcaaatctg cattggctac catgccaaca acagcaccga gcaagttgac 120
acgatcatgg agaagaacgt aaccgtcact cacgctcaag acatcctgga gaagaaacac 180
aacggtaagc tttgcgactt ggatggagtc aagcccctca tacttagaga ttgtagtgta 240
gccggttggc tgctcggtaa cccaatgtgc gatgagttca tcaatgttcc cgaatggtca 300
tatatcgtcg aaaaagctaa tcctgtcaac gacctgtgct accccggtga tttcaatgac 360
tatgaagaac tgaagcacct gctctcccgc atcaaccatt tcgagaaaat ccagatcatt 420
cccaagagtt cctggtctag ccatgaggct agtctgggtg tctcatccgc ctgcccatat 480
cagggtaaaa gttctttctt taggaacgta gtatggttga taaagaaaaa ctctacatac 540
ccgaccatca agcgctctta caacaatacg aaccaagagg atctgcttgt cctttgggga 600
atccatcatc ctaatgatgc tgccgaacag actaagctct accaaaaccc taccacttat 660
atttccgtcg gcacctctac tctgaaccag cgccttgtgc ccaggatcgc tacgagatca 720
aaagtcaacg gccaatcggg ccgcatggaa ttcttctgga cgatcctgaa gcctaatgac 780
gctatcaact tcgagtcaaa tggaaacttt atcgctcccg agtacgctta caagatcgtc 840
aagaagggcg actccacgat tatgaagtca gagttggagt acggcaactg caacacaaag 900
tgccaaactc ctatgggcgc tataaattct tcaatgccgt tccacaacat ccatccgctc 960
acgatcggtg agtgcccgaa atatgtaaag tcgaatcgtc tcgtactggc gacaggcctg 1020
agaaatagtc cgcaacgtga acgtcgtcgc aagaagagag gactgtttgg tgccattgca 1080
ggctttattg agggcggctg gcagggcatg gttgacggat ggtacggcta ccaccattca 1140
aacgagcagg gatctggcta cgccgctgac aaagaaagca cccaaaaggc cattgatgga 1200
gtgacgaata aggtgaattc gatcatcgac aaaatgaaca cgcaattcga agcagtgggt 1260
cgcgaattca ataacctgga gcgccgtatc gagaatctga acaagaagat ggaagacggc 1320
tttttggatg tctggacata taacgctgaa ttgctggtcc tcatggaaaa cgagcgtacc 1380
cttgatttcc acgacagcaa cgttaagaac ctctacgaca aggtcaggct ccagctcagg 1440
gataacgcca aggaattggg aaacggatgc ttcgagttct accacaaatg cgacaacgag 1500
tgcatggagt cagtcaggaa tggtacctac gactacccgc aatattctga ggaggctcgc 1560
ttgaagcgtg aggaaatatc gggtgttaaa ttggagagta ttggaatcta ccagcaccat 1620
caccaccatc actaataggc atgc 1644

<210> 147
<211> 909
<212> DNA
<213> Artificial Sequence

<220>
<223> HA1-1 IND

<400> 147
atgaaattct tagtcaacgt tgcccttgtt tttatggtcg tatacatttc ttacatctat 60
gcggagaaaa cccataacgg taagttgtgc gaccttgacg gtgtaaagcc cctgatcctc 120
cgtgactgca gtgttgctgg ttggcttttg ggcaacccca tgtgtgacga atttatcaac 180
gtgcctgaat ggtcatacat tgtagagaag gccaacccca cgaacgatct ctgttatccc 240
ggcagcttca atgactatga ggaacttaag caccttctgt cacgtatcaa ccacttcgaa 300
aagatccaga tcatcccgaa gagctcctgg agcgaccacg aagccagttc gggtgtgtct 360
tccgcttgcc cctacctcgg tagcccttcc ttcttccgta acgtagtgtg gctgatcaag 420
aagaatagca cttaccctac aatcaaaaag tcgtataaca atactaacca agaggatctg 480
cttgtactct ggggaattca tcatcccaac gacgcggcgg agcagaccag gttgtaccag 540
aaccccacca cttacatctc catcggtacg tccacactga atcagcgtct ggtccccaag 600
atcgcaacca ggtccaaggt taacggtcag tccggtcgta tggagttctt ctggaccatc 660
ctgaagccca acgacgccat caacttcgag tccaacggta atttcattgc tccggagtac 720

gcctacaaga tagttaagaa gggtgattca gcgatcatga agtcggaact tgagtatggc 780
aactgcaaca ctaaatgcca aactccaatg ggcgctatca actccagtat gccattccat 840
aacatccacc cattgacaat cggtgaatgt cccaagtacg tgaagcacca ccatcaccat 900
cactaatag 909

<210> 148
<211> 99
<212> DNA
<213> Artificial Sequence

<220>
<223> FOR PRIMER

<400> 148
aggcagatct atgaaattct tagtcaacgt tgcccttgtt tttatggtcg tatacatttc 60
ttacatctat gcggagaaaa cccataacgg taagttgtg 99

<210> 149
<211> 63
<212> DNA
<213> Artificial Sequence

<220>
<223> REV PRIMER

<400> 149
agtcgcatgc ctattaatgg tggtgatgat ggtgcttcac gtacttggga cattcaccga 60
ttg 63

<210> 150
<211> 1644
<212> DNA
<213> Artificial Sequence

<220>
<223> HA0s IND

<400> 150
agatctatga agttcctggt caatgtagcc ttggtattta tggtagtcta tatctcgtac 60
atttacgcag accagatttg tattggatat cacgctaaca acagcacaga gcaggtagat 120
actattatgg agaaaaatgt taccgtcact cacgcccagg acatcctgga gaaaacccat 180
aacggtaagt tgtgcgacct tgacggtgta aagcccctga tcctccgtga ctgcagtgtt 240
gctggttggc ttttgggcaa ccccatgtgt gacgaattta tcaacgtgcc tgaatggtca 300
tacattgtag agaaggccaa ccccacgaac gatctctgtt atcccggcag cttcaatgac 360
tatgaggaac ttaagcacct tctgtcacgt atcaaccact tcgaaaagat ccagatcatc 420
ccgaagagct cctggagcga ccacgaagcc agttcgggtg tgtcttccgc ttgcccctac 480
ctcggtagcc cttccttctt ccgtaacgta gtgtggctga tcaagaagaa tagcacttac 540
cctacaatca aaaagtcgta taacaatact aaccaagagg atctgcttgt actctgggga 600
attcatcatc ccaacgacgc ggcggagcag accaggttgt accagaaccc caccacttac 660
atctccatcg gtacgtccac actgaatcag cgtctggtcc ccaagatcgc aaccaggtcc 720
aaggttaacg gtcagtccgg tcgtatggag ttcttctgga ccatcctgaa gcccaacgac 780
gccatcaact tcgagtccaa cggtaatttc attgctccgg agtacgccta caagatagtt 840
aagaagggtg attcagcgat catgaagtcg gaacttgagt atggcaactg caacactaaa 900
tgccaaactc caatgggcgc tatcaactcc agtatgccat tccataacat ccacccattg 960
acaatcggtg aatgtcccaa gtacgtgaag agcaacaggt tggtattggc caccggtttg 1020
agaaacagcc cccagagaga gtcgcgtcgt aaaaagcgcg gcttgttcgg agccatcgct 1080
ggcttcatag agggtggttg gcagggaatg gtcgatggtt ggtatggtta tcatcattcc 1140
aacgagcagg gaagtggtta cgccgccgac aaagaatcga cccagaaggc tattgacggc 1200
gtcacaaaca aagtaaactc tatcattgat aagatgaaca cccagttcga ggctgtaggt 1260

agagaattca acaacctcga aagacgtatt gagaacctga acaagaaaat ggaggatggc 1320
ttcctggacg tgtggaccta caatgctgag ctgttggtcc ttatggagaa cgagcgtacc 1380
ctcgatttcc atgactcaaa cgtgaagaac ctgtatgaca aggtgcgttt gcaactgagg 1440
gacaacgcaa aggagcttgg aaacggttgt ttcgaatttt atcataagtg cgacaatgag 1500
tgtatggagt cgattagaaa tggcacgtac aactaccctc aatacagcga agaagctcgt 1560
ctcaaacgtg aggaaatcag cggcgtcaag ctcgaatcaa tcggtaccta tcagcaccac 1620
catcaccatc actaataggc atgc 1644

<210> 151
<211> 811
<212> PRT
<213> Artificial Sequence

<220>
<223> STF2.HA1-1His (PR8)

<400> 151
Met Lys Phe Leu Val Asn Val Ala Leu Val Phe Met Val Val Tyr Ile
1 5 10 15
Ser Tyr Ile Tyr Ala Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu
20 25 30
Leu Thr Gln Asn Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala
35 40 45
Ile Glu Arg Leu Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp
50 55 60
Ala Ala Gly Gln Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly
65 70 75 80
Leu Thr Gln Ala Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln
85 90 95
Thr Thr Glu Gly Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val
100 105 110
Arg Glu Leu Ala Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp
115 120 125
Leu Asp Ser Ile Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp
130 135 140
Arg Val Ser Gly Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln
145 150 155 160
Asp Asn Thr Leu Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile
165 170 175
Asp Ile Asp Leu Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser
180 185 190
Leu Asn Val Gln Lys Ala Tyr Asp Val Lys Asp Thr Ala Val Thr Thr
195 200 205
Lys Ala Tyr Ala Asn Asn Gly Thr Thr Leu Asp Val Ser Gly Leu Asp
210 215 220
Asp Ala Ala Ile Lys Ala Ala Thr Gly Gly Thr Asn Gly Thr Ala Ser
225 230 235 240
Val Thr Gly Gly Ala Val Lys Phe Asp Ala Asp Asn Asn Lys Tyr Phe
245 250 255
Val Thr Ile Gly Gly Phe Thr Gly Ala Asp Ala Ala Lys Asn Gly Asp
260 265 270
Tyr Glu Val Asn Val Ala Thr Asp Gly Thr Val Thr Leu Ala Ala Gly
275 280 285
Ala Thr Lys Thr Thr Met Pro Ala Gly Ala Thr Thr Lys Thr Glu Val
290 295 300
Gln Glu Leu Lys Asp Thr Pro Ala Val Val Ser Ala Asp Ala Lys Asn
305 310 315 320
Ala Leu Ile Ala Gly Gly Val Asp Ala Thr Asp Ala Asn Gly Ala Glu
325 330 335

```
Leu Val Lys Met Ser Tyr Thr Asp Lys Asn Gly Lys Thr Ile Glu Gly
            340                 345                 350
Gly Tyr Ala Leu Lys Ala Gly Asp Lys Tyr Tyr Ala Ala Asp Tyr Asp
        355                 360                 365
Glu Ala Thr Gly Ala Ile Lys Ala Lys Thr Thr Ser Tyr Thr Ala Ala
    370                 375                 380
Asp Gly Thr Thr Lys Thr Ala Ala Asn Gln Leu Gly Gly Val Asp Gly
385                 390                 395                 400
Lys Thr Glu Val Val Thr Ile Asp Gly Lys Thr Tyr Asn Ala Ser Lys
                405                 410                 415
Ala Ala Gly His Asp Phe Lys Ala Gln Pro Glu Leu Ala Glu Ala Ala
            420                 425                 430
Ala Lys Thr Thr Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala
        435                 440                 445
Gln Val Asp Ala Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe
    450                 455                 460
Asn Ser Ala Ile Thr Asn Leu Gly Asn Thr Val Asn Asn Leu Ser Glu
465                 470                 475                 480
Ala Arg Ser Arg Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn
                485                 490                 495
Met Ser Arg Ala Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala
            500                 505                 510
Gln Ala Asn Gln Val Pro Gln Asn Val Leu Ser Leu Leu Ala Ser Gly
        515                 520                 525
Ser Gly Ser Gly Ser Ser His Asn Gly Lys Leu Cys Arg Leu Lys Gly
    530                 535                 540
Ile Ala Pro Leu Gln Leu Gly Lys Cys Asn Ile Ala Gly Trp Leu Leu
545                 550                 555                 560
Gly Asn Pro Glu Cys Asp Pro Leu Leu Pro Val Arg Ser Trp Ser Tyr
                565                 570                 575
Ile Val Glu Thr Pro Asn Ser Glu Asn Gly Ile Cys Tyr Pro Gly Asp
            580                 585                 590
Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser
        595                 600                 605
Phe Glu Arg Phe Glu Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His
    610                 615                 620
Asn Thr Asn Gly Val Thr Ala Ala Cys Ser His Glu Gly Lys Ser Ser
625                 630                 635                 640
Phe Tyr Arg Asn Leu Leu Trp Leu Thr Glu Lys Glu Gly Ser Tyr Pro
                645                 650                 655
Lys Leu Lys Asn Ser Tyr Val Asn Lys Lys Gly Lys Glu Val Leu Val
            660                 665                 670
Leu Trp Gly Ile His His Pro Pro Asn Ser Lys Glu Gln Gln Asn Leu
        675                 680                 685
Tyr Gln Asn Glu Asn Ala Tyr Val Ser Val Val Thr Ser Asn Tyr Asn
    690                 695                 700
Arg Arg Phe Thr Pro Glu Ile Ala Glu Arg Pro Lys Val Arg Asp Gln
705                 710                 715                 720
Ala Gly Arg Met Asn Tyr Tyr Trp Thr Leu Leu Lys Pro Gly Asp Thr
                725                 730                 735
Ile Ile Phe Glu Ala Asn Gly Asn Leu Ile Ala Pro Met Tyr Ala Phe
            740                 745                 750
Ala Leu Ser Arg Gly Phe Gly Ser Gly Ile Ile Thr Ser Asn Ala Ser
        755                 760                 765
Met His Glu Cys Asn Thr Lys Cys Gln Thr Pro Leu Gly Ala Ile Asn
    770                 775                 780
Ser Ser Leu Pro Tyr Gln Asn Ile His Pro Val Thr Ile Gly Glu Cys
785                 790                 795                 800
Pro Lys Tyr Val Arg His His His His His His
```

805 810

<210> 152
<211> 811
<212> PRT
<213> Artificial Sequence

<220>
<223> ngSTF2.HA1-1His (PR8)

<400> 152
Met Lys Phe Leu Val Asn Val Ala Leu Val Phe Met Val Val Tyr Ile
1 5 10 15
Ser Tyr Ile Tyr Ala Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu
20 25 30
Leu Thr Gln Asn Asn Leu Gln Lys Ser Gln Ser Ala Leu Gly Thr Ala
35 40 45
Ile Glu Arg Leu Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp
50 55 60
Ala Ala Gly Gln Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly
65 70 75 80
Leu Thr Gln Ala Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln
85 90 95
Thr Thr Glu Gly Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val
100 105 110
Arg Glu Leu Ala Val Gln Ser Ala Gln Ser Thr Asn Ser Gln Ser Asp
115 120 125
Leu Asp Ser Ile Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp
130 135 140
Arg Val Ser Gly Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln
145 150 155 160
Asp Asn Thr Leu Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile
165 170 175
Asp Ile Asp Leu Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser
180 185 190
Leu Asn Val Gln Lys Ala Tyr Asp Val Lys Asp Thr Ala Val Thr Thr
195 200 205
Lys Ala Tyr Ala Asn Gln Gly Thr Thr Leu Asp Val Ser Gly Leu Asp
210 215 220
Asp Ala Ala Ile Lys Ala Ala Thr Gly Gly Thr Gln Gly Thr Ala Ser
225 230 235 240
Val Thr Gly Gly Ala Val Lys Phe Asp Ala Asp Asn Asn Lys Tyr Phe
245 250 255
Val Thr Ile Gly Gly Phe Thr Gly Ala Asp Ala Ala Lys Asn Gly Asp
260 265 270
Tyr Glu Val Asn Val Ala Thr Asp Gly Thr Val Thr Leu Ala Ala Gly
275 280 285
Ala Thr Lys Thr Thr Met Pro Ala Gly Ala Thr Thr Lys Thr Glu Val
290 295 300
Gln Glu Leu Lys Asp Thr Pro Ala Val Val Ser Ala Asp Ala Lys Asn
305 310 315 320
Ala Leu Ile Ala Gly Gly Val Asp Ala Thr Asp Ala Asn Gly Ala Glu
325 330 335
Leu Val Lys Met Ser Tyr Thr Asp Lys Asn Gly Lys Thr Ile Glu Gly
340 345 350
Gly Tyr Ala Leu Lys Ala Gly Asp Lys Tyr Tyr Ala Ala Asp Tyr Asp
355 360 365
Glu Ala Thr Gly Ala Ile Lys Ala Lys Thr Thr Ser Tyr Thr Ala Ala

```
    370                 375                 380
Asp Gly Thr Thr Lys Thr Ala Ala Asn Gln Leu Gly Gly Val Asp Gly
385                 390                 395                 400
Lys Thr Glu Val Val Thr Ile Asp Gly Lys Thr Tyr Gln Ala Ser Lys
                405                 410                 415
Ala Ala Gly His Asp Phe Lys Ala Gln Pro Glu Leu Ala Glu Ala Ala
            420                 425                 430
Ala Lys Thr Thr Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala
        435                 440                 445
Gln Val Asp Ala Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe
    450                 455                 460
Asn Ser Ala Ile Thr Asn Leu Gly Asn Thr Val Asn Gln Leu Ser Glu
465                 470                 475                 480
Ala Arg Ser Arg Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Gln
                485                 490                 495
Met Ser Arg Ala Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala
            500                 505                 510
Gln Ala Asn Gln Val Pro Gln Asn Val Leu Ser Leu Leu Ala Ser Gly
        515                 520                 525
Ser Gly Ser Gly Ser Ser His Asn Gly Lys Leu Cys Arg Leu Lys Gly
    530                 535                 540
Ile Ala Pro Leu Gln Leu Gly Lys Cys Asn Ile Ala Gly Trp Leu Leu
545                 550                 555                 560
Gly Asn Pro Glu Cys Asp Pro Leu Leu Pro Val Arg Ser Trp Ser Tyr
                565                 570                 575
Ile Val Glu Thr Pro Asn Ser Glu Asn Gly Ile Cys Tyr Pro Gly Asp
            580                 585                 590
Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser
        595                 600                 605
Phe Glu Arg Phe Glu Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His
    610                 615                 620
Asn Thr Asn Gly Val Thr Ala Ala Cys Ser His Glu Gly Lys Ser Ser
625                 630                 635                 640
Phe Tyr Arg Asn Leu Leu Trp Leu Thr Glu Lys Glu Gly Ser Tyr Pro
                645                 650                 655
Lys Leu Lys Asn Ser Tyr Val Asn Lys Lys Gly Lys Glu Val Leu Val
            660                 665                 670
Leu Trp Gly Ile His His Pro Pro Asn Ser Lys Glu Gln Gln Asn Leu
        675                 680                 685
Tyr Gln Asn Glu Asn Ala Tyr Val Ser Val Val Thr Ser Asn Tyr Asn
    690                 695                 700
Arg Arg Phe Thr Pro Glu Ile Ala Glu Arg Pro Lys Val Arg Asp Gln
705                 710                 715                 720
Ala Gly Arg Met Asn Tyr Tyr Trp Thr Leu Leu Lys Pro Gly Asp Thr
                725                 730                 735
Ile Ile Phe Glu Ala Asn Gly Asn Leu Ile Ala Pro Met Tyr Ala Phe
            740                 745                 750
Ala Leu Ser Arg Gly Phe Gly Ser Gly Ile Ile Thr Ser Gln Ala Ser
        755                 760                 765
Met His Glu Cys Asn Thr Lys Cys Gln Thr Pro Leu Gly Ala Ile Gln
    770                 775                 780
Ser Ser Leu Pro Tyr Gln Asn Ile His Pro Val Thr Ile Gly Glu Cys
785                 790                 795                 800
Pro Lys Tyr Val Arg His His His His His His
                805                 810
```

<210> 153
<211> 761

<212> PRT
<213> Artificial Sequence

<220>
<223> STF2.HA1-2His (PR8)

<400> 153

```
Met Lys Phe Leu Val Asn Val Ala Leu Val Phe Met Val Val Tyr Ile
 1               5                  10                  15
Ser Tyr Ile Tyr Ala Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu
            20                  25                  30
Leu Thr Gln Asn Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala
        35                  40                  45
Ile Glu Arg Leu Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp
    50                  55                  60
Ala Ala Gly Gln Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly
65                  70                  75                  80
Leu Thr Gln Ala Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln
                85                  90                  95
Thr Thr Glu Gly Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val
            100                 105                 110
Arg Glu Leu Ala Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp
        115                 120                 125
Leu Asp Ser Ile Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp
    130                 135                 140
Arg Val Ser Gly Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln
145                 150                 155                 160
Asp Asn Thr Leu Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile
                165                 170                 175
Asp Ile Asp Leu Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser
            180                 185                 190
Leu Asn Val Gln Lys Ala Tyr Asp Val Lys Asp Thr Ala Val Thr Thr
        195                 200                 205
Lys Ala Tyr Ala Asn Asn Gly Thr Thr Leu Asp Val Ser Gly Leu Asp
    210                 215                 220
Asp Ala Ala Ile Lys Ala Ala Thr Gly Gly Thr Asn Gly Thr Ala Ser
225                 230                 235                 240
Val Thr Gly Gly Ala Val Lys Phe Asp Ala Asp Asn Asn Lys Tyr Phe
                245                 250                 255
Val Thr Ile Gly Gly Phe Thr Gly Ala Asp Ala Ala Lys Asn Gly Asp
            260                 265                 270
Tyr Glu Val Asn Val Ala Thr Asp Gly Thr Val Thr Leu Ala Ala Gly
        275                 280                 285
Ala Thr Lys Thr Thr Met Pro Ala Gly Ala Thr Thr Lys Thr Glu Val
    290                 295                 300
Gln Glu Leu Lys Asp Thr Pro Ala Val Val Ser Ala Asp Ala Lys Asn
305                 310                 315                 320
Ala Leu Ile Ala Gly Gly Val Asp Ala Thr Asp Ala Asn Gly Ala Glu
                325                 330                 335
Leu Val Lys Met Ser Tyr Thr Asp Lys Asn Gly Lys Thr Ile Glu Gly
            340                 345                 350
Gly Tyr Ala Leu Lys Ala Gly Asp Lys Tyr Tyr Ala Ala Asp Tyr Asp
        355                 360                 365
Glu Ala Thr Gly Ala Ile Lys Ala Lys Thr Thr Ser Tyr Thr Ala Ala
    370                 375                 380
Asp Gly Thr Thr Lys Thr Ala Ala Asn Gln Leu Gly Gly Val Asp Gly
385                 390                 395                 400
Lys Thr Glu Val Val Thr Ile Asp Gly Lys Thr Tyr Asn Ala Ser Lys
                405                 410                 415
```

```
Ala Ala Gly His Asp Phe Lys Ala Gln Pro Glu Leu Ala Glu Ala Ala
            420                 425                 430
Ala Lys Thr Thr Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala
        435                 440                 445
Gln Val Asp Ala Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe
    450                 455                 460
Asn Ser Ala Ile Thr Asn Leu Gly Asn Thr Val Asn Asn Leu Ser Glu
465                 470                 475                 480
Ala Arg Ser Arg Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn
                485                 490                 495
Met Ser Arg Ala Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala
            500                 505                 510
Gln Ala Asn Gln Val Pro Gln Asn Val Leu Ser Leu Leu Ala Ser Gly
        515                 520                 525
Ser Gly Ser Gly Ser Lys Gly Ile Ala Pro Gln Leu Gly Lys Cys Asn
    530                 535                 540
Ile Ala Gly Trp Leu Leu Gly Asn Pro Glu Cys Asp Pro Leu Leu Pro
545                 550                 555                 560
Val Arg Ser Trp Ser Tyr Ile Val Glu Thr Pro Asn Ser Glu Asn Gly
                565                 570                 575
Ile Cys Tyr Pro Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln
            580                 585                 590
Leu Ser Ser Val Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Glu
        595                 600                 605
Ser Ser Trp Pro Asn His Asn Thr Asn Gly Val Thr Ala Ala Cys Ser
    610                 615                 620
His Glu Gly Lys Ser Ser Phe Tyr Arg Asn Leu Leu Trp Leu Thr Glu
625                 630                 635                 640
Lys Glu Gly Ser Tyr Pro Lys Leu Lys Asn Ser Tyr Val Asn Lys Lys
                645                 650                 655
Gly Lys Glu Val Leu Val Leu Trp Gly Ile His His Pro Pro Asn Ser
            660                 665                 670
Lys Glu Gln Gln Asn Leu Tyr Gln Asn Glu Asn Ala Tyr Val Ser Val
        675                 680                 685
Val Thr Ser Asn Tyr Asn Arg Arg Phe Thr Pro Glu Ile Ala Glu Arg
    690                 695                 700
Pro Lys Val Arg Asp Gln Ala Gly Arg Met Asn Tyr Tyr Trp Thr Leu
705                 710                 715                 720
Leu Lys Pro Gly Asp Thr Ile Ile Phe Glu Ala Asn Gly Asn Leu Ile
                725                 730                 735
Ala Pro Met Tyr Ala Phe Ala Leu Ser Arg Gly Phe Gly Ser Gly Ile
            740                 745                 750
Ile Thr Ser His His His His His His
        755                 760


<210> 154
<211> 762
<212> PRT
<213> Artificial Sequence

<220>
<223> ngSTF2.HA1-2His (PR8)

<400> 154
Met Lys Phe Leu Val Asn Val Ala Leu Val Phe Met Val Val Tyr Ile
 1                  5                  10                  15
Ser Tyr Ile Tyr Ala Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu
            20                  25                  30
```

```
Leu Thr Gln Asn Asn Leu Gln Lys Ser Gln Ser Ala Leu Gly Thr Ala
        35                  40                  45
Ile Glu Arg Leu Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp
    50                  55                  60
Ala Ala Gly Gln Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly
65                  70                  75                  80
Leu Thr Gln Ala Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln
                85                  90                  95
Thr Thr Glu Gly Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val
            100                 105                 110
Arg Glu Leu Ala Val Gln Ser Ala Gln Ser Thr Asn Ser Gln Ser Asp
        115                 120                 125
Leu Asp Ser Ile Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp
    130                 135                 140
Arg Val Ser Gly Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln
145                 150                 155                 160
Asp Asn Thr Leu Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile
                165                 170                 175
Asp Ile Asp Leu Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser
            180                 185                 190
Leu Asn Val Gln Lys Ala Tyr Asp Val Lys Asp Thr Ala Val Thr Thr
        195                 200                 205
Lys Ala Tyr Ala Asn Gln Gly Thr Thr Leu Asp Val Ser Gly Leu Asp
    210                 215                 220
Asp Ala Ala Ile Lys Ala Ala Thr Gly Gly Thr Gln Gly Thr Ala Ser
225                 230                 235                 240
Val Thr Gly Gly Ala Val Lys Phe Asp Ala Asp Asn Asn Lys Tyr Phe
                245                 250                 255
Val Thr Ile Gly Gly Phe Thr Gly Ala Asp Ala Ala Lys Asn Gly Asp
            260                 265                 270
Tyr Glu Val Asn Val Ala Thr Asp Gly Thr Val Thr Leu Ala Ala Gly
        275                 280                 285
Ala Thr Lys Thr Thr Met Pro Ala Gly Ala Thr Thr Lys Thr Glu Val
    290                 295                 300
Gln Glu Leu Lys Asp Thr Pro Ala Val Val Ser Ala Asp Ala Lys Asn
305                 310                 315                 320
Ala Leu Ile Ala Gly Gly Val Asp Ala Thr Asp Ala Asn Gly Ala Glu
                325                 330                 335
Leu Val Lys Met Ser Tyr Thr Asp Lys Asn Gly Lys Thr Ile Glu Gly
            340                 345                 350
Gly Tyr Ala Leu Lys Ala Gly Asp Lys Tyr Tyr Ala Ala Asp Tyr Asp
        355                 360                 365
Glu Ala Thr Gly Ala Ile Lys Ala Lys Thr Thr Ser Tyr Thr Ala Ala
    370                 375                 380
Asp Gly Thr Thr Lys Thr Ala Ala Asn Gln Leu Gly Gly Val Asp Gly
385                 390                 395                 400
Lys Thr Glu Val Val Thr Ile Asp Gly Lys Thr Tyr Gln Ala Ser Lys
                405                 410                 415
Ala Ala Gly His Asp Phe Lys Ala Gln Pro Glu Leu Ala Glu Ala Ala
            420                 425                 430
Ala Lys Thr Thr Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala
        435                 440                 445
Gln Val Asp Ala Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe
    450                 455                 460
Asn Ser Ala Ile Thr Asn Leu Gly Asn Thr Val Asn Gln Leu Ser Glu
465                 470                 475                 480
Ala Arg Ser Arg Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Gln
                485                 490                 495
Met Ser Arg Ala Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala
```

```
            500                 505                 510
Gln Ala Asn Gln Val Pro Gln Asn Val Leu Ser Leu Leu Ala Ser Gly
        515                 520                 525
Ser Gly Ser Gly Ser Lys Gly Ile Ala Pro Leu Gln Leu Gly Lys Cys
    530                 535                 540
Asn Ile Ala Gly Trp Leu Leu Gly Asn Pro Glu Cys Asp Pro Leu Leu
545                 550                 555                 560
Pro Val Arg Ser Trp Ser Tyr Ile Val Glu Thr Pro Asn Ser Glu Asn
                565                 570                 575
Gly Ile Cys Tyr Pro Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu
            580                 585                 590
Gln Leu Ser Ser Val Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys
        595                 600                 605
Glu Ser Ser Trp Pro Asn His Asn Thr Asn Gly Val Thr Ala Ala Cys
    610                 615                 620
Ser His Glu Gly Lys Ser Ser Phe Tyr Arg Asn Leu Leu Trp Leu Thr
625                 630                 635                 640
Glu Lys Glu Gly Ser Tyr Pro Lys Leu Lys Asn Ser Tyr Val Asn Lys
                645                 650                 655
Lys Gly Lys Glu Val Leu Val Leu Trp Gly Ile His His Pro Pro Asn
            660                 665                 670
Ser Lys Glu Gln Gln Asn Leu Tyr Gln Asn Glu Asn Ala Tyr Val Ser
        675                 680                 685
Val Val Thr Ser Asn Tyr Asn Arg Arg Phe Thr Pro Glu Ile Ala Glu
    690                 695                 700
Arg Pro Lys Val Arg Asp Gln Ala Gly Arg Met Asn Tyr Tyr Trp Thr
705                 710                 715                 720
Leu Leu Lys Pro Gly Asp Thr Ile Ile Phe Glu Ala Asn Gly Asn Leu
                725                 730                 735
Ile Ala Pro Met Tyr Ala Phe Ala Leu Ser Arg Gly Phe Gly Ser Gly
            740                 745                 750
Ile Ile Thr Ser His His His His His His
        755                 760


<210> 155
<211> 762
<212> PRT
<213> Artificial Sequence

<220>
<223> STF2.HA1-2mutHis (PR8)

<400> 155
Met Lys Phe Leu Val Asn Val Ala Leu Val Phe Met Val Val Tyr Ile
 1                5                  10                  15
Ser Tyr Ile Tyr Ala Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu
            20                  25                  30
Leu Thr Gln Asn Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala
        35                  40                  45
Ile Glu Arg Leu Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp
    50                  55                  60
Ala Ala Gly Gln Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly
65                  70                  75                  80
Leu Thr Gln Ala Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln
                85                  90                  95
Thr Thr Glu Gly Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val
            100                 105                 110
Arg Glu Leu Ala Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp
```

```
        115                 120                 125
Leu Asp Ser Ile Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp
    130                 135                 140
Arg Val Ser Gly Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln
145                 150                 155                 160
Asp Asn Thr Leu Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile
                165                 170                 175
Asp Ile Asp Leu Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser
            180                 185                 190
Leu Asn Val Gln Lys Ala Tyr Asp Val Lys Asp Thr Ala Val Thr Thr
        195                 200                 205
Lys Ala Tyr Ala Asn Asn Gly Thr Thr Leu Asp Val Ser Gly Leu Asp
    210                 215                 220
Asp Ala Ala Ile Lys Ala Ala Thr Gly Gly Thr Asn Gly Thr Ala Ser
225                 230                 235                 240
Val Thr Gly Gly Ala Val Lys Phe Asp Ala Asp Asn Asn Lys Tyr Phe
                245                 250                 255
Val Thr Ile Gly Gly Phe Thr Gly Ala Asp Ala Ala Lys Asn Gly Asp
            260                 265                 270
Tyr Glu Val Asn Val Ala Thr Asp Gly Thr Val Thr Leu Ala Ala Gly
        275                 280                 285
Ala Thr Lys Thr Thr Met Pro Ala Gly Ala Thr Thr Lys Thr Glu Val
    290                 295                 300
Gln Glu Leu Lys Asp Thr Pro Ala Val Val Ser Ala Asp Ala Lys Asn
305                 310                 315                 320
Ala Leu Ile Ala Gly Gly Val Asp Ala Thr Asp Ala Asn Gly Ala Glu
                325                 330                 335
Leu Val Lys Met Ser Tyr Thr Asp Lys Asn Gly Lys Thr Ile Glu Gly
            340                 345                 350
Gly Tyr Ala Leu Lys Ala Gly Asp Lys Tyr Tyr Ala Ala Asp Tyr Asp
        355                 360                 365
Glu Ala Thr Gly Ala Ile Lys Ala Lys Thr Thr Ser Tyr Thr Ala Ala
    370                 375                 380
Asp Gly Thr Thr Lys Thr Ala Ala Asn Gln Leu Gly Gly Val Asp Gly
385                 390                 395                 400
Lys Thr Glu Val Val Thr Ile Asp Gly Lys Thr Tyr Asn Ala Ser Lys
                405                 410                 415
Ala Ala Gly His Asp Phe Lys Ala Gln Pro Glu Leu Ala Glu Ala Ala
            420                 425                 430
Ala Lys Thr Thr Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala
        435                 440                 445
Gln Val Asp Ala Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe
    450                 455                 460
Asn Ser Ala Ile Thr Asn Leu Gly Asn Thr Val Asn Asn Leu Ser Glu
465                 470                 475                 480
Ala Arg Ser Arg Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn
                485                 490                 495
Met Ser Arg Ala Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala
            500                 505                 510
Gln Ala Asn Gln Val Pro Gln Asn Val Leu Ser Leu Leu Ala Ser Gly
        515                 520                 525
Ser Gly Ser Gly Ser Lys Gly Ala Ala Pro Leu Gln Leu Gly Lys Cys
    530                 535                 540
Asn Ile Ala Gly Trp Leu Leu Gly Asn Pro Glu Cys Asp Pro Leu Leu
545                 550                 555                 560
Pro Val Arg Ser Trp Ser Asp Ile Ala Glu Thr Pro Asn Ser Glu Asn
                565                 570                 575
Gly Ile Cys Tyr Pro Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu
            580                 585                 590
```

```
Gln Leu Ser Ser Val Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys
        595                 600                 605
Glu Ser Ser Trp Pro Asn His Asn Thr Asn Gly Val Thr Ala Ala Cys
    610                 615                 620
Ser His Glu Gly Lys Ser Ser Phe Tyr Arg Asn Leu Leu Trp Leu Thr
625                 630                 635                 640
Glu Lys Glu Gly Ser Tyr Pro Lys Leu Lys Asn Ser Tyr Val Asn Lys
                645                 650                 655
Lys Gly Lys Glu Val Leu Val Leu Trp Gly Ile His His Pro Pro Asn
            660                 665                 670
Ser Lys Glu Gln Gln Asn Leu Tyr Gln Asn Glu Asn Ala Tyr Val Ser
        675                 680                 685
Val Val Thr Ser Asn Tyr Asn Arg Arg Phe Thr Pro Glu Ile Ala Glu
    690                 695                 700
Arg Pro Lys Val Arg Asp Gln Ala Gly Arg Met Asn Tyr Tyr Trp Thr
705                 710                 715                 720
Leu Leu Lys Pro Gly Asp Thr Ile Ile Phe Glu Ala Asn Gly Asn Leu
                725                 730                 735
Ile Ala Pro Met Tyr Ala Phe Ala Leu Ser Arg Gly Phe Gly Ser Gly
            740                 745                 750
Ile Ile Thr Ser His His His His His His
        755                 760


<210> 156
<211> 761
<212> PRT
<213> Artificial Sequence

<220>
<223> ngSTF2.HA1-2mutHis (PR8)

<400> 156
Met Lys Phe Leu Val Asn Val Ala Leu Val Phe Met Val Val Tyr Ile
 1                  5                  10                  15
Ser Tyr Ile Tyr Ala Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu
            20                  25                  30
Leu Thr Gln Asn Asn Leu Gln Lys Ser Gln Ser Ala Leu Gly Thr Ala
        35                  40                  45
Ile Glu Arg Leu Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp
    50                  55                  60
Ala Ala Gly Gln Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly
65                  70                  75                  80
Leu Thr Gln Ala Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln
                85                  90                  95
Thr Thr Glu Gly Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val
            100                 105                 110
Arg Glu Leu Ala Val Gln Ser Ala Gln Ser Thr Asn Ser Gln Ser Asp
        115                 120                 125
Leu Asp Ser Ile Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp
    130                 135                 140
Arg Val Ser Gly Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln
145                 150                 155                 160
Asp Asn Thr Leu Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile
                165                 170                 175
Asp Ile Asp Leu Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser
            180                 185                 190
Leu Asn Val Gln Lys Ala Tyr Asp Val Lys Asp Thr Ala Val Thr Thr
        195                 200                 205
```

```
Lys Ala Tyr Ala Asn Gln Gly Thr Thr Leu Asp Val Ser Gly Leu Asp
    210                 215                 220
Asp Ala Ala Ile Lys Ala Ala Thr Gly Gly Thr Gln Gly Thr Ala Ser
225                 230                 235                 240
Val Thr Gly Gly Ala Val Lys Phe Asp Ala Asp Asn Asn Lys Tyr Phe
                245                 250                 255
Val Thr Ile Gly Gly Phe Thr Gly Ala Asp Ala Ala Lys Asn Gly Asp
            260                 265                 270
Tyr Glu Val Asn Val Ala Thr Asp Gly Thr Val Thr Leu Ala Ala Gly
        275                 280                 285
Ala Thr Lys Thr Thr Met Pro Ala Gly Ala Thr Thr Lys Thr Glu Val
    290                 295                 300
Gln Glu Leu Lys Asp Thr Pro Ala Val Val Ser Ala Asp Ala Lys Asn
305                 310                 315                 320
Ala Leu Ile Ala Gly Gly Val Asp Ala Thr Asp Ala Asn Gly Ala Glu
                325                 330                 335
Leu Val Lys Met Ser Tyr Thr Asp Lys Asn Gly Lys Thr Ile Glu Gly
            340                 345                 350
Gly Tyr Ala Leu Lys Ala Gly Asp Lys Tyr Tyr Ala Ala Asp Tyr Asp
        355                 360                 365
Glu Ala Thr Gly Ala Ile Lys Ala Lys Thr Thr Ser Tyr Thr Ala Ala
    370                 375                 380
Asp Gly Thr Thr Lys Thr Ala Ala Asn Gln Leu Gly Gly Val Asp Gly
385                 390                 395                 400
Lys Thr Glu Val Val Thr Ile Asp Gly Lys Thr Tyr Gln Ala Ser Lys
                405                 410                 415
Ala Ala Gly His Asp Phe Lys Ala Gln Pro Glu Leu Ala Glu Ala Ala
            420                 425                 430
Ala Lys Thr Thr Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala
        435                 440                 445
Gln Val Asp Ala Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe
    450                 455                 460
Asn Ser Ala Ile Thr Asn Leu Gly Asn Thr Val Asn Gln Leu Ser Glu
465                 470                 475                 480
Ala Arg Ser Arg Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Gln
                485                 490                 495
Met Ser Arg Ala Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala
            500                 505                 510
Gln Ala Asn Gln Val Pro Gln Asn Val Leu Ser Leu Leu Ala Ser Gly
        515                 520                 525
Ser Gly Ser Gly Ser Lys Gly Ala Ala Pro Gln Leu Gly Lys Cys Asn
    530                 535                 540
Ile Ala Gly Trp Leu Leu Gly Asn Pro Glu Cys Asp Pro Leu Leu Pro
545                 550                 555                 560
Val Arg Ser Trp Ser Asp Ile Ala Glu Thr Pro Asn Ser Glu Asn Gly
                565                 570                 575
Ile Cys Tyr Pro Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln
            580                 585                 590
Leu Ser Ser Val Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Glu
        595                 600                 605
Ser Ser Trp Pro Asn His Asn Thr Asn Gly Val Thr Ala Ala Cys Ser
    610                 615                 620
His Glu Gly Lys Ser Ser Phe Tyr Arg Asn Leu Leu Trp Leu Thr Glu
625                 630                 635                 640
Lys Glu Gly Ser Tyr Pro Lys Leu Lys Asn Ser Tyr Val Asn Lys Lys
                645                 650                 655
Gly Lys Glu Val Leu Val Leu Trp Gly Ile His His Pro Pro Asn Ser
            660                 665                 670
Lys Glu Gln Gln Asn Leu Tyr Gln Asn Glu Asn Ala Tyr Val Ser Val
```

```
        675                 680                 685
Val Thr Ser Asn Tyr Asn Arg Arg Phe Thr Pro Glu Ile Ala Glu Arg
    690                 695                 700
Pro Lys Val Arg Asp Gln Ala Gly Arg Met Asn Tyr Tyr Trp Thr Leu
705                 710                 715                 720
Leu Lys Pro Gly Asp Thr Ile Ile Phe Glu Ala Asn Gly Asn Leu Ile
                725                 730                 735
Ala Pro Met Tyr Ala Phe Ala Leu Ser Arg Gly Phe Gly Ser Gly Ile
            740                 745                 750
Ile Thr Ser His His His His His His
        755                 760


<210> 157
<211> 714
<212> PRT
<213> Artificial Sequence

<220>
<223> STF2.HA1-3His (PR8)

<400> 157
Met Lys Phe Leu Val Asn Val Ala Leu Val Phe Met Val Val Tyr Ile
 1               5                  10                  15
Ser Tyr Ile Tyr Ala Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu
            20                  25                  30
Leu Thr Gln Asn Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala
        35                  40                  45
Ile Glu Arg Leu Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp
    50                  55                  60
Ala Ala Gly Gln Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly
65                  70                  75                  80
Leu Thr Gln Ala Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln
                85                  90                  95
Thr Thr Glu Gly Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val
            100                 105                 110
Arg Glu Leu Ala Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp
        115                 120                 125
Leu Asp Ser Ile Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp
    130                 135                 140
Arg Val Ser Gly Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln
145                 150                 155                 160
Asp Asn Thr Leu Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile
                165                 170                 175
Asp Ile Asp Leu Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser
            180                 185                 190
Leu Asn Val Gln Lys Ala Tyr Asp Val Lys Asp Thr Ala Val Thr Thr
        195                 200                 205
Lys Ala Tyr Ala Asn Asn Gly Thr Thr Leu Asp Val Ser Gly Leu Asp
    210                 215                 220
Asp Ala Ala Ile Lys Ala Ala Thr Gly Gly Thr Asn Gly Thr Ala Ser
225                 230                 235                 240
Val Thr Gly Gly Ala Val Lys Phe Asp Ala Asp Asn Asn Lys Tyr Phe
                245                 250                 255
Val Thr Ile Gly Gly Phe Thr Gly Ala Asp Ala Ala Lys Asn Gly Asp
            260                 265                 270
Tyr Glu Val Asn Val Ala Thr Asp Gly Thr Val Thr Leu Ala Ala Gly
        275                 280                 285
Ala Thr Lys Thr Thr Met Pro Ala Gly Ala Thr Thr Lys Thr Glu Val
```

```
    290                 295                 300
Gln Glu Leu Lys Asp Thr Pro Ala Val Val Ser Ala Asp Ala Lys Asn
305                 310                 315                 320
Ala Leu Ile Ala Gly Gly Val Asp Ala Thr Asp Ala Asn Gly Ala Glu
                325                 330                 335
Leu Val Lys Met Ser Tyr Thr Asp Lys Asn Gly Lys Thr Ile Glu Gly
            340                 345                 350
Gly Tyr Ala Leu Lys Ala Gly Asp Lys Tyr Tyr Ala Ala Asp Tyr Asp
        355                 360                 365
Glu Ala Thr Gly Ala Ile Lys Ala Lys Thr Thr Ser Tyr Thr Ala Ala
    370                 375                 380
Asp Gly Thr Thr Lys Thr Ala Ala Asn Gln Leu Gly Gly Val Asp Gly
385                 390                 395                 400
Lys Thr Glu Val Val Thr Ile Asp Gly Lys Thr Tyr Asn Ala Ser Lys
                405                 410                 415
Ala Ala Gly His Asp Phe Lys Ala Gln Pro Glu Leu Ala Glu Ala Ala
            420                 425                 430
Ala Lys Thr Thr Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala
        435                 440                 445
Gln Val Asp Ala Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe
    450                 455                 460
Asn Ser Ala Ile Thr Asn Leu Gly Asn Thr Val Asn Asn Leu Ser Glu
465                 470                 475                 480
Ala Arg Ser Arg Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn
                485                 490                 495
Met Ser Arg Ala Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala
            500                 505                 510
Gln Ala Asn Gln Val Pro Gln Asn Val Leu Ser Leu Leu Ala Ser Gly
        515                 520                 525
Ser Gly Ser Gly Ser Asn Ser Glu Asn Gly Cys Tyr Pro Gly Asp Phe
    530                 535                 540
Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe
545                 550                 555                 560
Glu Arg Phe Glu Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His Asn
                565                 570                 575
Thr Asn Gly Val Thr Ala Ala Cys Ser His Glu Gly Lys Ser Ser Phe
            580                 585                 590
Tyr Arg Asn Leu Leu Trp Leu Thr Glu Lys Glu Gly Ser Tyr Pro Lys
        595                 600                 605
Leu Lys Asn Ser Tyr Val Asn Lys Lys Gly Lys Glu Val Leu Val Leu
    610                 615                 620
Trp Gly Ile His His Pro Pro Asn Ser Lys Glu Gln Gln Asn Leu Tyr
625                 630                 635                 640
Gln Asn Glu Asn Ala Tyr Val Ser Val Val Thr Ser Asn Tyr Asn Arg
                645                 650                 655
Arg Phe Thr Pro Glu Ile Ala Glu Arg Pro Lys Val Arg Asp Gln Ala
            660                 665                 670
Gly Arg Met Asn Tyr Tyr Trp Thr Leu Leu Lys Pro Gly Asp Thr Ile
        675                 680                 685
Ile Phe Glu Ala Asn Gly Asn Leu Ile Ala Pro Met Tyr Ala Phe Ala
    690                 695                 700
Leu Ser Arg Gly His His His His His His
705                 710
```

<210> 158
<211> 715
<212> PRT
<213> Artificial Sequence

<220>
<223> STF2.HA1-3mutHis (PR8)

<400> 158

```
Met Lys Phe Leu Val Asn Val Ala Leu Val Phe Met Val Val Tyr Ile
 1               5                  10                  15
Ser Tyr Ile Tyr Ala Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu
            20                  25                  30
Leu Thr Gln Asn Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala
        35                  40                  45
Ile Glu Arg Leu Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp
    50                  55                  60
Ala Ala Gly Gln Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly
65                  70                  75                  80
Leu Thr Gln Ala Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln
                85                  90                  95
Thr Thr Glu Gly Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val
            100                 105                 110
Arg Glu Leu Ala Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp
        115                 120                 125
Leu Asp Ser Ile Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp
    130                 135                 140
Arg Val Ser Gly Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln
145                 150                 155                 160
Asp Asn Thr Leu Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile
                165                 170                 175
Asp Ile Asp Leu Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser
            180                 185                 190
Leu Asn Val Gln Lys Ala Tyr Asp Val Lys Asp Thr Ala Val Thr Thr
        195                 200                 205
Lys Ala Tyr Ala Asn Asn Gly Thr Thr Leu Asp Val Ser Gly Leu Asp
    210                 215                 220
Asp Ala Ala Ile Lys Ala Ala Thr Gly Gly Thr Asn Gly Thr Ala Ser
225                 230                 235                 240
Val Thr Gly Gly Ala Val Lys Phe Asp Ala Asp Asn Asn Lys Tyr Phe
                245                 250                 255
Val Thr Ile Gly Gly Phe Thr Gly Ala Asp Ala Ala Lys Asn Gly Asp
            260                 265                 270
Tyr Glu Val Asn Val Ala Thr Asp Gly Thr Val Thr Leu Ala Ala Gly
        275                 280                 285
Ala Thr Lys Thr Thr Met Pro Ala Gly Ala Thr Thr Lys Thr Glu Val
    290                 295                 300
Gln Glu Leu Lys Asp Thr Pro Ala Val Val Ser Ala Asp Ala Lys Asn
305                 310                 315                 320
Ala Leu Ile Ala Gly Gly Val Asp Ala Thr Asp Ala Asn Gly Ala Glu
                325                 330                 335
Leu Val Lys Met Ser Tyr Thr Asp Lys Asn Gly Lys Thr Ile Glu Gly
            340                 345                 350
Gly Tyr Ala Leu Lys Ala Gly Asp Lys Tyr Tyr Ala Ala Asp Tyr Asp
        355                 360                 365
Glu Ala Thr Gly Ala Ile Lys Ala Lys Thr Thr Ser Tyr Thr Ala Ala
    370                 375                 380
Asp Gly Thr Thr Lys Thr Ala Ala Asn Gln Leu Gly Gly Val Asp Gly
385                 390                 395                 400
Lys Thr Glu Val Val Thr Ile Asp Gly Lys Thr Tyr Asn Ala Ser Lys
                405                 410                 415
Ala Ala Gly His Asp Phe Lys Ala Gln Pro Glu Leu Ala Glu Ala Ala
            420                 425                 430
```

```
Ala Lys Thr Thr Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala
        435                 440                 445
Gln Val Asp Ala Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe
    450                 455                 460
Asn Ser Ala Ile Thr Asn Leu Gly Asn Thr Val Asn Asn Leu Ser Glu
465                 470                 475                 480
Ala Arg Ser Arg Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn
                485                 490                 495
Met Ser Arg Ala Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala
            500                 505                 510
Gln Ala Asn Gln Val Pro Gln Asn Val Leu Ser Leu Leu Ala Ser Gly
        515                 520                 525
Ser Gly Ser Gly Ser Asn Ser Glu Asn Glu Ile Cys Tyr Pro Gly Asp
    530                 535                 540
Phe Ile Asp Lys Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser
545                 550                 555                 560
Phe Glu Arg Phe Glu Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His
                565                 570                 575
Asn Thr Asn Gly Val Thr Ala Ala Cys Ser His Glu Gly Lys Ser Ser
            580                 585                 590
Phe Tyr Arg Asn Leu Leu Trp Leu Thr Glu Lys Glu Gly Ser Tyr Pro
        595                 600                 605
Lys Leu Lys Asn Ser Tyr Val Asn Lys Lys Gly Lys Glu Val Leu Val
    610                 615                 620
Leu Trp Gly Ile His His Pro Pro Asn Ser Lys Glu Gln Gln Asn Leu
625                 630                 635                 640
Tyr Gln Asn Glu Asn Ala Tyr Val Ser Val Val Thr Ser Asn Tyr Asn
                645                 650                 655
Arg Arg Phe Thr Pro Glu Ile Ala Glu Arg Pro Lys Val Arg Asp Gln
            660                 665                 670
Ala Gly Arg Met Asn Tyr Tyr Trp Thr Leu Leu Lys Pro Gly Asp Thr
        675                 680                 685
Ile Ile Phe Glu Ala Asn Gly Asn Leu Ile Ala Pro Met Tyr Ala Ala
    690                 695                 700
Ala Leu Ser Arg Gly His His His His His His
705                 710                 715


<210> 159
<211> 729
<212> PRT
<213> Artificial Sequence

<220>
<223> STF2.HA1-2 (IND)

<400> 159
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
 1                5                  10                  15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
            20                  25                  30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
        35                  40                  45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
    50                  55                  60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65                  70                  75                  80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
                85                  90                  95
```

```
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
            100                 105                 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
        115                 120                 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
    130                 135                 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145                 150                 155                 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val Gln
                165                 170                 175
Lys Ala Tyr Asp Val Lys Asp Thr Ala Val Thr Thr Lys Ala Tyr Ala
            180                 185                 190
Asn Asn Gly Thr Thr Leu Asp Val Ser Gly Leu Asp Asp Ala Ala Ile
        195                 200                 205
Lys Ala Ala Thr Gly Gly Thr Asn Gly Thr Ala Ser Val Thr Gly Gly
    210                 215                 220
Ala Val Lys Phe Asp Ala Asp Asn Asn Lys Tyr Phe Val Thr Ile Gly
225                 230                 235                 240
Gly Phe Thr Gly Ala Asp Ala Ala Lys Asn Gly Asp Tyr Glu Val Asn
                245                 250                 255
Val Ala Thr Asp Gly Thr Val Thr Leu Ala Ala Gly Ala Thr Lys Thr
            260                 265                 270
Thr Met Pro Ala Gly Ala Thr Thr Lys Thr Glu Val Gln Glu Leu Lys
        275                 280                 285
Asp Thr Pro Ala Val Val Ser Ala Asp Ala Lys Asn Ala Leu Ile Ala
    290                 295                 300
Gly Gly Val Asp Ala Thr Asp Ala Asn Gly Ala Glu Leu Val Lys Met
305                 310                 315                 320
Ser Tyr Thr Asp Lys Asn Gly Lys Thr Ile Glu Gly Gly Tyr Ala Leu
                325                 330                 335
Lys Ala Gly Asp Lys Tyr Tyr Ala Ala Asp Tyr Asp Glu Ala Thr Gly
            340                 345                 350
Ala Ile Lys Ala Lys Thr Thr Ser Tyr Thr Ala Ala Asp Gly Thr Thr
        355                 360                 365
Lys Thr Ala Ala Asn Gln Leu Gly Gly Val Asp Gly Lys Thr Glu Val
    370                 375                 380
Val Thr Ile Asp Gly Lys Thr Tyr Asn Ala Ser Lys Ala Ala Gly His
385                 390                 395                 400
Asp Phe Lys Ala Gln Pro Glu Leu Ala Glu Ala Ala Ala Lys Thr Thr
                405                 410                 415
Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala
            420                 425                 430
Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile
        435                 440                 445
Thr Asn Leu Gly Asn Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg
    450                 455                 460
Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala
465                 470                 475                 480
Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln
                485                 490                 495
Val Pro Gln Asn Val Leu Ser Leu Leu Ala Gly Val Lys Pro Leu Ile
            500                 505                 510
Leu Arg Asp Cys Ser Val Ala Gly Trp Leu Leu Gly Asn Pro Met Cys
        515                 520                 525
Asp Glu Phe Ile Asn Val Pro Glu Trp Ser Tyr Ile Val Glu Lys Ala
    530                 535                 540
Asn Pro Thr Asn Asp Leu Cys Tyr Pro Gly Ser Phe Asn Asp Tyr Glu
545                 550                 555                 560
Glu Leu Lys His Leu Leu Ser Arg Ile Asn His Phe Glu Lys Ile Gln
```

```
                565                 570                 575
Ile Ile Pro Lys Ser Ser Trp Ser Asp His Glu Ala Ser Ser Gly Val
            580                 585                 590
Ser Ser Ala Cys Pro Tyr Leu Gly Ser Pro Ser Phe Phe Arg Asn Val
        595                 600                 605
Val Trp Leu Ile Lys Lys Asn Ser Thr Tyr Pro Thr Ile Lys Lys Ser
    610                 615                 620
Tyr Asn Asn Thr Asn Gln Glu Asp Leu Leu Val Leu Trp Gly Ile His
625                 630                 635                 640
His Pro Asn Asp Ala Ala Glu Gln Thr Arg Leu Tyr Gln Asn Pro Thr
                645                 650                 655
Thr Tyr Ile Ser Ile Gly Thr Ser Thr Leu Asn Gln Arg Leu Val Pro
            660                 665                 670
Lys Ile Ala Thr Arg Ser Lys Val Asn Gly Gln Ser Gly Arg Met Glu
        675                 680                 685
Phe Phe Trp Thr Ile Leu Lys Pro Asn Asp Ala Ile Asn Phe Glu Ser
    690                 695                 700
Asn Gly Asn Phe Ile Ala Pro Glu Tyr Ala Tyr Lys Ile Val Lys Lys
705                 710                 715                 720
Gly Asp Ser Ala Ile Met Lys Ser Glu
                725
```

<210> 160
<211> 2433
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2delta.HA0shis

<400> 160

```
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcatgg agcgccggtg 540
gatcctgcta gcccatggac cgaaaacccg ctgcagaaaa ttgatgccgc gctggcgcag 600
gtggatgcgc tgcgctctga tctgggtgcg gtacaaaacc gtttcaactc tgctatcacc 660
aacctgggca ataccgtaaa caatctgtct gaagcgcgta gccgtatcga agattccgac 720
tacgcgaccg aagtttccaa catgtctcgc gcgcagattt tgcagcaggc cggtacttcc 780
gttctggcgc aggctaacca ggtcccgcag aacgtgctgt ctctgttacg tgaattctct 840
agatatccag cacagtggcg gccgctcgac acaatatgta taggctacca tgcgaacaat 900
tcaaccgaca ctgttgacac agtactcgag aagaatgtga cagtgacaca ctctgttaac 960
ctgctcgaag acagccacaa cggaaaacta tgtagattaa aaggaatagc cccactacaa 1020
ttggggaaat gtaacatcgc cggatggctt ttgggaaacc cagaatgcga cccactgctt 1080
ccagtgagat catggtccta cattgtagaa acaccaaact ctgagaatgg aatatgttat 1140
ccaggagatt tcatcgacta tgaggagctg agggagcaat tgagctcagt gtcatcattc 1200
gaaagattcg aaatatttcc caaagaaagc tcatggccca accacaacac aaacggagta 1260
acggcagcat gctcccatga ggggaaaagc agtttttaca gaaatttgct atggctgacg 1320
gagaaggagg gctcataccc aaagctgaaa aattcttatg tgaacaaaaa agggaaagaa 1380
gtccttgtac tgtggggtat tcatcacccg cctaacagta aggaacaaca gaatctctat 1440
cagaatgaaa atgcttatgt ctctgtagtg acttcaaatt ataacaggag atttaccccg 1500
gaaatagcag aaagacccaa agtaagagat caagctggga ggatgaacta ttactggacc 1560
ttgctaaaac ccggagacac aataatattt gaggcaaatg gaaatctaat agcaccaatg 1620
```

tatgctttcg cactgagtag aggctttggg tccggcatca tcacctcaaa cgcatcaatg 1680
catgagtgta acacgaagtg tcaaacaccc ctgggagcta taaacagcag tctcccttac 1740
cagaatatac acccagtcac aataggagag tgcccaaaat acgtcaggag tgccaaattg 1800
aggatggtta caggactaag gaacattccg tccattcaat ccagaggtct atttggagcc 1860
attgccggtt ttattgaagg gggatggact ggaatgatag atggatggta tggttatcat 1920
catcagaatg aacagggatc aggctatgca gcggatcaaa aaagcacaca aaatgccatt 1980
aacgggatta caaacaaggt gaacactgtt atcgagaaaa tgaacattca attcacagct 2040
gtgggtaaag aattcaacaa attagaaaaa aggatggaaa atttaaataa aaaagttgat 2100
gatggatttc tggacatttg gacatataat gcagaattgt tagttctact ggaaaatgaa 2160
aggactctgg atttccatga ctcaaatgtg aagaatctgt atgagaaagt aaaaagccaa 2220
ttaaagaata atgccaaaga aatcggaaat ggatgttttg agttctacca caagtgtgac 2280
aatgaatgca tggaaagtgt aagaaatggg acttatgatt atcccaaata ttcagaagag 2340
tcaaagttga acagggaaaa ggtagatgga gtgaaattgg aatcaatggg gatctatcag 2400
acgcgtaccg gtcatcatca ccatcaccat tga 2433

<210> 161
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> FOR PRIMER

<400> 161
attctccggc ggccgctcga cacaatatgt ataggctacc 40

<210> 162
<211> 60
<212> DNA
<213> Artificial Sequence

<220>
<223> REV PRIMER

<400> 162
agtcttgcgg ccgcctatta atggtgatgg tgatgatgct gatagatccc cattgattcc 60

<210> 163
<211> 1533
<212> DNA
<213> Artificial Sequence

<220>
<223> HA0s PR8 template

<400> 163
gacacaatat gtataggcta ccatgcgaac aattcaaccg acactgttga cacagtactc 60
gagaagaatg tgacagtgac acactctgtt aacctgctcg aagacagcca caacggaaaa 120
ctatgtagat taaaaggaat agccccacta caattgggga aatgtaacat cgccggatgg 180
cttttgggaa acccagaatg cgacccactg cttccagtga gatcatggtc ctacattgta 240
gaaacaccaa actctgagaa tggaatatgt tatccaggag atttcatcga ctatgaggag 300
ctgagggagc aattgagctc agtgtcatca ttcgaaagat tcgaaatatt tcccaaagaa 360
agctcatggc ccaaccacaa cacaaacgga gtaacggcag catgctccca tgaggggaaa 420
agcagttttt acagaaattt gctatggctg acggagaagg agggctcata cccaaagctg 480
aaaaattctt atgtgaacaa aaaagggaaa gaagtccttg tactgtgggg tattcatcac 540
ccgcctaaca gtaaggaaca acagaatctc tatcagaatg aaaatgctta tgtctctgta 600
gtgacttcaa attataacag gagatttacc ccggaaatag cagaaagacc caaagtaaga 660
gatcaagctg ggaggatgaa ctattactgg accttgctaa aacccggaga cacaataata 720

tttgaggcaa atggaaatct aatagcacca atgtatgctt tcgcactgag tagaggcttt 780
gggtccggca tcatcacctc aaacgcatca atgcatgagt gtaacacgaa gtgtcaaaca 840
cccctgggag ctataaacag cagtctccct taccagaata tacacccagt cacaatagga 900
gagtgcccaa aatacgtcag gagtgccaaa ttgaggatgg ttacaggact aaggaacatt 960
ccgtccattc aatccagagg tctatttgga gccattgccg gttttattga agggggatgg 1020
actggaatga tagatggatg gtatggttat catcatcaga atgaacaggg atcaggctat 1080
gcagcggatc aaaaaagcac acaaaatgcc attaacggga ttacaaacaa ggtgaacact 1140
gttatcgaga aaatgaacat tcaattcaca gctgtgggta aagaattcaa caaattagaa 1200
aaaaggatgg aaaatttaaa taaaaaagtt gatgatggat ttctggacat ttggacatat 1260
aatgcagaat tgttagttct actggaaaat gaaaggactc tggatttcca tgactcaaat 1320
gtgaagaatc tgtatgagaa agtaaaaagc caattaaaga ataatgccaa agaaatcgga 1380
aatggatgtt ttgagttcta ccacaagtgt gacaatgaat gcatggaaag tgtaagaaat 1440
gggacttatg attatcccaa atattcagaa gagtcaaagt tgaacaggga aaaggtagat 1500
ggagtgaaat tggaatcaat ggggatctat cag 1533

<210> 164
<211> 2403
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2delta.HA0s

<400> 164
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcatgg agcgccggtg 540
gatcctgcta gcccatggac cgaaaacccg ctgcagaaaa ttgatgccgc gctggcgcag 600
gtggatgcgc tgcgctctga tctgggtgcg gtacaaaacc gtttcaactc tgctatcacc 660
aacctgggca ataccgtaaa caatctgtct gaagcgcgta gccgtatcga agattccgac 720
tacgcgaccg aagtttccaa catgtctcgc gcgcagattt tgcagcaggc cggtacttcc 780
gttctggcgc aggctaacca ggtcccgcag aacgtgctgt ctctgttacg tgaattctct 840
agatatccag cacagtggcg gccgctcgac acaatatgta taggctacca tgcgaacaat 900
tcaaccgaca ctgttgacac agtactcgag aagaatgtga cagtgacaca ctctgttaac 960
ctgctcgaag acagccacaa cggaaaacta tgtagattaa aaggaatagc cccactacaa 1020
ttggggaaat gtaacatcgc cggatggctt ttgggaaacc cagaatgcga cccactgctt 1080
ccagtgagat catggtccta cattgtagaa acaccaaact ctgagaatgg aatatgttat 1140
ccaggagatt tcatcgacta tgaggagctg agggagcaat tgagctcagt gtcatcattc 1200
gaaagattcg aaatatttcc caaagaaagc tcatggccca accacaacac aaacggagta 1260
acggcagcat gctcccatga ggggaaaagc agtttttaca gaaatttgct atggctgacg 1320
gagaaggagg gctcataccc aaagctgaaa aattcttatg tgaacaaaaa agggaaagaa 1380
gtccttgtac tgtggggtat tcatcacccg cctaacagta aggaacaaca gaatctctat 1440
cagaatgaaa atgcttatgt ctctgtagtg acttcaaatt ataacaggag atttaccccg 1500
gaaatagcag aaagacccaa agtaagagat caagctggga ggatgaacta ttactggacc 1560
ttgctaaaac ccggagacac aataatattt gaggcaaatg gaaatctaat agcaccaatg 1620
tatgctttcg cactgagtag aggctttggg tccggcatca tcacctcaaa cgcatcaatg 1680
catgagtgta acacgaagtg tcaaacaccc ctgggagcta taaacagcag tctcccttac 1740
cagaatatac acccagtcac aataggagag tgcccaaaat acgtcaggag tgccaaattg 1800
aggatggtta caggactaag gaacattccg tccattcaat ccagaggtct atttggagcc 1860
attgccggtt ttattgaagg gggatggact ggaatgatag atggatggta tggttatcat 1920
catcagaatg aacagggatc aggctatgca gcggatcaaa aaagcacaca aaatgccatt 1980
aacgggatta caaacaaggt gaacactgtt atcgagaaaa tgaacattca attcacagct 2040
gtgggtaaag aattcaacaa attagaaaaa aggatggaaa atttaaataa aaaagttgat 2100

gatggatttc tggacatttg gacatataat gcagaattgt tagttctact ggaaaatgaa 2160
aggactctgg atttccatga ctcaaatgtg aagaatctgt atgagaaagt aaaaagccaa 2220
ttaaagaata atgccaaaga aatcggaaat ggatgttttg agttctacca caagtgtgac 2280
aatgaatgca tggaaagtgt aagaaatggg acttatgatt atcccaaata ttcagaagag 2340
tcaaagttga acagggaaaa ggtagatgga gtgaaattgg aatcaatggg gatctatcag 2400
tag 2403

<210> 165
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> FOR PRIMER

<400> 165
attctccggc ggccgctcga cacaatatgt ataggctacc 40

<210> 166
<211> 60
<212> DNA
<213> Artificial Sequence

<220>
<223> REV PRIMER

<400> 166
agtcttgcgg ccgcctatta atggtgatgg tgatgatgct gatagatccc cattgattcc 60

<210> 167
<211> 2403
<212> DNA
<213> Artificial Sequence

<220>
<223> HA0s.STF2delta

<400> 167
gacacaatat gtataggcta ccatgcgaac aattcaaccg acactgttga cacagtactc 60
gagaagaatg tgacagtgac acactctgtt aacctgctcg aagacagcca caacggaaaa 120
ctatgtagat taaaaggaat agccccacta caattgggga aatgtaacat cgccggatgg 180
cttttgggaa acccagaatg cgacccactg cttccagtga gatcatggtc ctacattgta 240
gaaacaccaa actctgagaa tggaatatgt tatccaggag atttcatcga ctatgaggag 300
ctgagggagc aattgagctc agtgtcatca ttcgaaagat tcgaaatatt tcccaaagaa 360
agctcatggc ccaaccacaa cacaaacgga gtaacggcag catgctccca tgaggggaaa 420
agcagttttt acagaaattt gctatggctg acggagaagg agggctcata cccaaagctg 480
aaaaattctt atgtgaacaa aaaagggaaa gaagtccttg tactgtgggg tattcatcac 540
ccgcctaaca gtaaggaaca acagaatctc tatcagaatg aaaatgctta tgtctctgta 600
gtgacttcaa attataacag gagatttacc ccggaaatag cagaaagacc caaagtaaga 660
gatcaagctg ggaggatgaa ctattactgg accttgctaa aacccggaga cacaataata 720
tttgaggcaa atggaaatct aatagcacca atgtatgctt tcgcactgag tagaggcttt 780
gggtccggca tcatcacctc aaacgcatca atgcatgagt gtaacacgaa gtgtcaaaca 840
cccctgggag ctataaacag cagtctccct taccagaata tacacccagt cacaatagga 900
gagtgcccaa aatacgtcag gagtgccaag ttgaggatgg ttacaggact aaggaacatt 960
ccgtccattc aatccagagg tctatttgga gccattgccg gttttattga aggggatgg 1020
actggaatga tagatggatg gtatggttat catcatcaga atgaacaggg atcaggctat 1080
gcagcggatc aaaaaagcac acaaaatgcc attaacggga ttacaaacaa ggtgaacact 1140
gttatcgaga aaatgaacat tcaattcaca gctgtgggta aagaattcaa caaattagaa 1200

aaaaggatgg aaaatttaaa taaaaaagtt gatgatggat ttctggacat ttggacatat 1260
aatgcagaat tgttagttct actggaaaat gaaaggactc tggatttcca tgactcaaat 1320
gtgaagaatc tgtatgagaa agtaaaaagc caattaaaga ataatgccaa agaaatcgga 1380
aatggatgtt ttgagttcta ccacaagtgt gacaatgaat gcatggaaag tgtaagaaat 1440
gggacttatg attatcccaa atattcagaa gagtcaaagt tgaacaggga aaaggtagat 1500
ggagtgaaat tggaatcaat ggggatctat caggaattct ctagatatcc agcacagtgg 1560
cggccgctcg cacaagtaat caacactaac agtctgtcgc tgctgaccca gaataacctg 1620
aacaaatccc agtccgcact gggcaccgct atcgagcgtc tgtcttctgg tctgcgtatc 1680
aacagcgcga aagacgatgc ggcaggtcag gcgattgcta accgtttcac cgcgaacatc 1740
aaaggtctga ctcaggcttc ccgtaacgct aacgacggta tctccattgc gcagaccact 1800
gaaggcgcgc tgaacgaaat caacaacaac ctgcagcgtg tgcgtgaact ggcggttcag 1860
tctgctaaca gcaccaactc ccagtctgac ctcgactcca tccaggctga aatcacccag 1920
cgcctgaacg aaatcgaccg tgtatccggc cagactcagt tcaacggcgt gaaagtcctg 1980
gcgcaggaca acaccctgac catccaggtt ggcgccaacg acggtgaaac tatcgatatc 2040
gatctgaagc agatcaactc tcagaccctg ggtctggact cactgaacgt gcatggagcg 2100
ccggtggatc ctgctagccc atggaccgaa aacccgctgc agaaaattga tgccgcgctg 2160
gcgcaggtgg atgcgctgcg ctctgatctg ggtgcggtac aaaaccgttt caactctgct 2220
atcaccaacc tgggcaatac cgtaaacaat ctgtctgaag cgcgtagccg tatcgaagat 2280
tccgactacg cgaccgaagt ttccaacatg tctcgcgcgc agattttgca gcaggccggt 2340
acttccgttc tggcgcaggc taaccaggtc ccgcagaacg tgctgtctct gttacgttaa 2400
tag 2403

<210> 168
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> FOR PRIMER

<400> 168
aggcaagatc tgacacaata tgtataggct acc 33

<210> 169
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> REV PRIMER

<400> 169
agtcagacgc gtctattaac gtaacagaga cagcac 36

<210> 170
<211> 1566
<212> DNA
<213> Artificial Sequence

<220>
<223> Ha0s.his

<400> 170
gacacaatat gtataggcta ccatgcgaac aattcaaccg acactgttga cacagtactc 60
gagaagaatg tgacagtgac acactctgtt aacctgctcg aagacagcca caacggaaaa 120
ctatgtagat taaaaggaat agccccacta caattgggga aatgtaacat cgccggatgg 180
cttttgggaa acccagaatg cgacccactg cttccagtga gatcatggtc ctacattgta 240
gaaacaccaa actctgagaa tggaatatgt tatccaggag atttcatcga ctatgaggag 300
ctgagggagc aattgagctc agtgtcatca ttcgaaagat tcgaaatatt tcccaaagaa 360

agctcatggc ccaaccacaa cacaaacgga gtaacggcag catgctccca tgaggggaaa 420
agcagttttt acagaaattt gctatggctg acggagaagg agggctcata cccaaagctg 480
aaaaattctt atgtgaacaa aaaagggaaa gaagtccttg tactgtgggg tattcatcac 540
ccgcctaaca gtaaggaaca acagaatctc tatcagaatg aaaatgctta tgtctctgta 600
gtgacttcaa attataacag gagatttacc ccggaaatag cagaaagacc caaagtaaga 660
gatcaagctg ggaggatgaa ctattactgg accttgctaa aacccggaga cacaataata 720
tttgaggcaa atggaaatct aatagcacca atgtatgctt tcgcactgag tagaggcttt 780
gggtccggca tcatcacctc aaacgcatca atgcatgagt gtaacacgaa gtgtcaaaca 840
cccctgggag ctataaacag cagtctccct taccagaata tacacccagt cacaatagga 900
gagtgcccaa aatacgtcag gagtgccaaa ttgaggatgg ttacaggact aaggaacatt 960
ccgtccattc aatccagagg tctatttgga gccattgccg gttttattga agggggatgg 1020
actggaatga tagatggatg gtatggttat catcatcaga atgaacaggg atcaggctat 1080
gcagcggatc aaaaaagcac acaaaatgcc attaacggga ttacaaacaa ggtgaacact 1140
gttatcgaga aaatgaacat tcaattcaca gctgtgggta aagaattcaa caaattagaa 1200
aaaaggatgg aaaatttaaa taaaaaagtt gatgatggat ttctggacat ttggacatat 1260
aatgcagaat tgttagttct actggaaaat gaaaggactc tggatttcca tgactcaaat 1320
gtgaagaatc tgtatgagaa agtaaaaagc caattaaaga ataatgccaa agaaatcgga 1380
aatggatgtt ttgagttcta ccacaagtgt gacaatgaat gcatggaaag tgtaagaaat 1440
gggacttatg attatcccaa atattcagaa gagtcaaagt tgaacaggga aaaggtagat 1500
ggagtgaaat tggaatcaat ggggatctat cagacgcgta ccggtcatca tcaccatcac 1560
cattga 1566

<210> 171
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> FOR PRIMER

<400> 171
attctccggc ggccgctcga cacaatatgt ataggctacc 40

<210> 172
<211> 60
<212> DNA
<213> Artificial Sequence

<220>
<223> HA0s

<400> 172
agtcttgcgg ccgcctatta atggtgatgg tgatgatgct gatagatccc cattgattcc 60

<210> 173
<211> 1533
<212> DNA
<213> Artificial Sequence

<220>
<223> HA0s

<400> 173
gacacaatat gtataggcta ccatgcgaac aattcaaccg acactgttga cacagtactc 60
gagaagaatg tgacagtgac acactctgtt aacctgctcg aagacagcca caacggaaaa 120
ctatgtagat taaaaggaat agccccacta caattgggga aatgtaacat cgccggatgg 180
cttttgggaa acccagaatg cgacccactg cttccagtga gatcatggtc ctacattgta 240
gaaacaccaa actctgagaa tggaatatgt tatccaggag atttcatcga ctatgaggag 300

ctgagggagc aattgagctc agtgtcatca ttcgaaagat tcgaaatatt tcccaaagaa 360
agctcatggc ccaaccacaa cacaaacgga gtaacggcag catgctccca tgaggggaaa 420
agcagttttt acagaaattt gctatggctg acggagaagg agggctcata cccaaagctg 480
aaaaattctt atgtgaacaa aaaagggaaa gaagtccttg tactgtgggg tattcatcac 540
ccgcctaaca gtaaggaaca acagaatctc tatcagaatg aaaatgctta tgtctctgta 600
gtgacttcaa attataacag gagatttacc ccggaaatag cagaaagacc caaagtaaga 660
gatcaagctg ggaggatgaa ctattactgg accttgctaa aacccggaga cacaataata 720
tttgaggcaa atggaaatct aatagcacca atgtatgctt tcgcactgag tagaggcttt 780
gggtccggca tcatcacctc aaacgcatca atgcatgagt gtaacacgaa gtgtcaaaca 840
cccctgggag ctataaacag cagtctccct taccagaata tacacccagt cacaatagga 900
gagtgcccaa aatacgtcag gagtgccaaa ttgaggatgg ttacaggact aaggaacatt 960
ccgtccattc aatccagagg tctatttgga gccattgccg gttttattga agggggatgg 1020
actggaatga tagatggatg gtatggttat catcatcaga atgaacaggg atcaggctat 1080
gcagcggatc aaaaaagcac acaaaatgcc attaacggga ttacaaacaa ggtgaacact 1140
gttatcgaga aaatgaacat tcaattcaca gctgtgggta aagaattcaa caaattagaa 1200
aaaaggatgg aaaatttaaa taaaaaagtt gatgatggat ttctggacat ttggacatat 1260
aatgcagaat tgttagttct actggaaaat gaaaggactc tggatttcca tgactcaaat 1320
gtgaagaatc tgtatgagaa agtaaaaagc caattaaaga ataatgccaa agaaatcgga 1380
aatggatgtt ttgagttcta ccacaagtgt gacaatgaat gcatggaaag tgtaagaaat 1440
gggacttatg attatcccaa atattcagaa gagtcaaagt tgaacaggga aaaggtagat 1500
ggagtgaaat tggaatcaat ggggatctat cag 1533

<210> 174
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> FOR PRIMER

<400> 174
aggcaagatc tgacacaata tgtataggct acc 33

<210> 175
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> REV PRIMER

<400> 175
agtcagacgc gtctattaac gtaacagaga cagcac 36

<210> 176
<211> 517
<212> PRT
<213> Artificial Sequence

<220>
<223> HA0sHis (PR8)

<400> 176
Asp Thr Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Asp Thr Val
1 5 10 15
Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His Ser Val Asn Leu
20 25 30
Leu Glu Asp Ser His Asn Gly Lys Leu Cys Arg Leu Lys Gly Ile Ala
35 40 45

```
Pro Leu Gln Leu Gly Lys Cys Asn Ile Ala Gly Trp Leu Leu Gly Asn
    50                  55                  60
Pro Glu Cys Asp Pro Leu Leu Pro Val Arg Ser Trp Ser Tyr Ile Val
65                  70                  75                  80
Glu Thr Pro Asn Ser Glu Asn Gly Ile Cys Tyr Pro Gly Asp Phe Ile
                85                  90                  95
Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe Glu
            100                 105                 110
Arg Phe Glu Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His Asn Thr
        115                 120                 125
Asn Gly Val Thr Ala Ala Cys Ser His Glu Gly Lys Ser Ser Phe Tyr
    130                 135                 140
Arg Asn Leu Leu Trp Leu Thr Glu Lys Glu Gly Ser Tyr Pro Lys Leu
145                 150                 155                 160
Lys Asn Ser Tyr Val Asn Lys Lys Gly Lys Glu Val Leu Val Leu Trp
                165                 170                 175
Gly Ile His His Pro Pro Asn Ser Lys Glu Gln Gln Asn Leu Tyr Gln
            180                 185                 190
Asn Glu Asn Ala Tyr Val Ser Val Val Thr Ser Asn Tyr Asn Arg Arg
        195                 200                 205
Phe Thr Pro Glu Ile Ala Glu Arg Pro Lys Val Arg Asp Gln Ala Gly
    210                 215                 220
Arg Met Asn Tyr Tyr Trp Thr Leu Leu Lys Pro Gly Asp Thr Ile Ile
225                 230                 235                 240
Phe Glu Ala Asn Gly Asn Leu Ile Ala Pro Met Tyr Ala Phe Ala Leu
                245                 250                 255
Ser Arg Gly Phe Gly Ser Gly Ile Ile Thr Ser Asn Ala Ser Met His
            260                 265                 270
Glu Cys Asn Thr Lys Cys Gln Thr Pro Leu Gly Ala Ile Asn Ser Ser
        275                 280                 285
Leu Pro Tyr Gln Asn Ile His Pro Val Thr Ile Gly Glu Cys Pro Lys
    290                 295                 300
Tyr Val Arg Ser Ala Lys Leu Arg Met Val Thr Gly Leu Arg Asn Ile
305                 310                 315                 320
Pro Ser Ile Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile
                325                 330                 335
Glu Gly Gly Trp Thr Gly Met Ile Asp Gly Trp Tyr Gly Tyr His His
            340                 345                 350
Gln Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Gln Lys Ser Thr Gln
        355                 360                 365
Asn Ala Ile Asn Gly Ile Thr Asn Lys Val Asn Thr Val Ile Glu Lys
    370                 375                 380
Met Asn Ile Gln Phe Thr Ala Val Gly Lys Glu Phe Asn Lys Leu Glu
385                 390                 395                 400
Lys Arg Met Glu Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Leu Asp
                405                 410                 415
Ile Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu Arg
            420                 425                 430
Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys Val
        435                 440                 445
Lys Ser Gln Leu Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys Phe
    450                 455                 460
Glu Phe Tyr His Lys Cys Asp Asn Glu Cys Met Glu Ser Val Arg Asn
465                 470                 475                 480
Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ser Lys Leu Asn Arg
                485                 490                 495
Glu Lys Val Asp Gly Val Lys Leu Glu Ser Met Gly Ile Tyr Gln His
            500                 505                 510
His His His His His
```

515

<210> 177
<211> 800
<212> PRT
<213> Artificial Sequence

<220>
<223> STF2delta.HA0s (PR8)

<400> 177

Met	Ala	Gln	Val	Ile	Asn	Thr	Asn	Ser	Leu	Ser	Leu	Leu	Thr	Gln	Asn
1				5				10						15	
Asn	Leu	Asn	Lys	Ser	Gln	Ser	Ala	Leu	Gly	Thr	Ala	Ile	Glu	Arg	Leu
			20					25					30		
Ser	Ser	Gly	Leu	Arg	Ile	Asn	Ser	Ala	Lys	Asp	Asp	Ala	Ala	Gly	Gln
		35					40					45			
Ala	Ile	Ala	Asn	Arg	Phe	Thr	Ala	Asn	Ile	Lys	Gly	Leu	Thr	Gln	Ala
	50					55					60				
Ser	Arg	Asn	Ala	Asn	Asp	Gly	Ile	Ser	Ile	Ala	Gln	Thr	Thr	Glu	Gly
65					70					75					80
Ala	Leu	Asn	Glu	Ile	Asn	Asn	Asn	Leu	Gln	Arg	Val	Arg	Glu	Leu	Ala
				85					90					95	
Val	Gln	Ser	Ala	Asn	Ser	Thr	Asn	Ser	Gln	Ser	Asp	Leu	Asp	Ser	Ile
			100					105					110		
Gln	Ala	Glu	Ile	Thr	Gln	Arg	Leu	Asn	Glu	Ile	Asp	Arg	Val	Ser	Gly
		115					120					125			
Gln	Thr	Gln	Phe	Asn	Gly	Val	Lys	Val	Leu	Ala	Gln	Asp	Asn	Thr	Leu
	130					135					140				
Thr	Ile	Gln	Val	Gly	Ala	Asn	Asp	Gly	Glu	Thr	Ile	Asp	Ile	Asp	Leu
145					150					155					160
Lys	Gln	Ile	Asn	Ser	Gln	Thr	Leu	Gly	Leu	Asp	Ser	Leu	Asn	Val	His
				165					170					175	
Gly	Ala	Pro	Val	Asp	Pro	Ala	Ser	Pro	Trp	Thr	Glu	Asn	Pro	Leu	Gln
			180					185					190		
Lys	Ile	Asp	Ala	Ala	Leu	Ala	Gln	Val	Asp	Ala	Leu	Arg	Ser	Asp	Leu
		195					200					205			
Gly	Ala	Val	Gln	Asn	Arg	Phe	Asn	Ser	Ala	Ile	Thr	Asn	Leu	Gly	Asn
	210					215					220				
Thr	Val	Asn	Asn	Leu	Ser	Glu	Ala	Arg	Ser	Arg	Ile	Glu	Asp	Ser	Asp
225					230					235					240
Tyr	Ala	Thr	Glu	Val	Ser	Asn	Met	Ser	Arg	Ala	Gln	Ile	Leu	Gln	Gln
				245					250					255	
Ala	Gly	Thr	Ser	Val	Leu	Ala	Gln	Ala	Asn	Gln	Val	Pro	Gln	Asn	Val
			260					265					270		
Leu	Ser	Leu	Leu	Arg	Glu	Phe	Ser	Arg	Tyr	Pro	Ala	Gln	Trp	Arg	Pro
		275					280					285			
Leu	Asp	Thr	Ile	Cys	Ile	Gly	Tyr	His	Ala	Asn	Asn	Ser	Thr	Asp	Thr
	290					295					300				
Val	Asp	Thr	Val	Leu	Glu	Lys	Asn	Val	Thr	Val	Thr	His	Ser	Val	Asn
305					310					315					320
Leu	Leu	Glu	Asp	Ser	His	Asn	Gly	Lys	Leu	Cys	Arg	Leu	Lys	Gly	Ile
				325					330					335	
Ala	Pro	Leu	Gln	Leu	Gly	Lys	Cys	Asn	Ile	Ala	Gly	Trp	Leu	Leu	Gly
			340					345					350		
Asn	Pro	Glu	Cys	Asp	Pro	Leu	Leu	Pro	Val	Arg	Ser	Trp	Ser	Tyr	Ile
		355					360					365			
Val	Glu	Thr	Pro	Asn	Ser	Glu	Asn	Gly	Ile	Cys	Tyr	Pro	Gly	Asp	Phe

```
    370                 375                 380
Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe
385                 390                 395                 400
Glu Arg Phe Glu Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His Asn
                405                 410                 415
Thr Asn Gly Val Thr Ala Ala Cys Ser His Glu Gly Lys Ser Ser Phe
            420                 425                 430
Tyr Arg Asn Leu Leu Trp Leu Thr Glu Lys Glu Gly Ser Tyr Pro Lys
        435                 440                 445
Leu Lys Asn Ser Tyr Val Asn Lys Lys Gly Lys Glu Val Leu Val Leu
    450                 455                 460
Trp Gly Ile His His Pro Pro Asn Ser Lys Glu Gln Gln Asn Leu Tyr
465                 470                 475                 480
Gln Asn Glu Asn Ala Tyr Val Ser Val Val Thr Ser Asn Tyr Asn Arg
                485                 490                 495
Arg Phe Thr Pro Glu Ile Ala Glu Arg Pro Lys Val Arg Asp Gln Ala
            500                 505                 510
Gly Arg Met Asn Tyr Tyr Trp Thr Leu Leu Lys Pro Gly Asp Thr Ile
        515                 520                 525
Ile Phe Glu Ala Asn Gly Asn Leu Ile Ala Pro Met Tyr Ala Phe Ala
    530                 535                 540
Leu Ser Arg Gly Phe Gly Ser Gly Ile Ile Thr Ser Asn Ala Ser Met
545                 550                 555                 560
His Glu Cys Asn Thr Lys Cys Gln Thr Pro Leu Gly Ala Ile Asn Ser
                565                 570                 575
Ser Leu Pro Tyr Gln Asn Ile His Pro Val Thr Ile Gly Glu Cys Pro
            580                 585                 590
Lys Tyr Val Arg Ser Ala Lys Leu Arg Met Val Thr Gly Leu Arg Asn
        595                 600                 605
Ile Pro Ser Ile Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe
    610                 615                 620
Ile Glu Gly Gly Trp Thr Gly Met Ile Asp Gly Trp Tyr Gly Tyr His
625                 630                 635                 640
His Gln Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Gln Lys Ser Thr
                645                 650                 655
Gln Asn Ala Ile Asn Gly Ile Thr Asn Lys Val Asn Thr Val Ile Glu
            660                 665                 670
Lys Met Asn Ile Gln Phe Thr Ala Val Gly Lys Glu Phe Asn Lys Leu
        675                 680                 685
Glu Lys Arg Met Glu Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Leu
    690                 695                 700
Asp Ile Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu
705                 710                 715                 720
Arg Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys
                725                 730                 735
Val Lys Ser Gln Leu Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys
            740                 745                 750
Phe Glu Phe Tyr His Lys Cys Asp Asn Glu Cys Met Glu Ser Val Arg
        755                 760                 765
Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ser Lys Leu Asn
    770                 775                 780
Arg Glu Lys Val Asp Gly Val Lys Leu Glu Ser Met Gly Ile Tyr Gln
785                 790                 795                 800
```

<210> 178
<211> 1518
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2

<400> 178
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcagaa agcgtatgat 540
gtgaaagata cagcagtaac aacgaaagct tatgccaata atggtactac actggatgta 600
tcgggtcttg atgatgcagc tattaaagcg gctacgggtg gtacgaatgg tacggcttct 660
gtaaccggtg gtgcggttaa atttgacgca gataataaca agtactttgt tactattggt 720
ggctttactg gtgctgatgc cgccaaaaat ggcgattatg aagttaacgt tgctactgac 780
ggtacagtaa cccttgcggc tggcgcaact aaaaccacaa tgcctgctgg tgcgacaact 840
aaaacagaag tacaggagtt aaaagataca ccggcagttg tttcagcaga tgctaaaaat 900
gccttaattg ctggcggcgt tgacgctacc gatgctaatg gcgctgagtt ggtcaaaatg 960
tcttataccg ataaaaatgg taagacaatt gaaggcggtt atgcgcttaa agctggcgat 1020
aagtattacg ccgcagatta cgatgaagcg acaggagcaa ttaaagctaa aactacaagt 1080
tatactgctg ctgacggcac taccaaaaca gcggctaacc aactgggtgg cgtagacggt 1140
aaaaccgaag tcgttactat cgacggtaaa acctacaatg ccagcaaagc cgctggtcat 1200
gatttcaaag cacaaccaga gctggcggaa gcagccgcta aaaccaccga aaacccgctg 1260
cagaaaattg atgccgcgct ggcgcaggtg gatgcgctgc gctctgatct gggtgcggta 1320
caaaaccgtt tcaactctgc tatcaccaac ctgggcaata ccgtaaacaa tctgtctgaa 1380
gcgcgtagcc gtatcgaaga ttccgactac gcgaccgaag tttccaacat gtctcgcgcg 1440
cagattctgc agcaggccgg tacttccgtt ctggcgcagg ctaaccaggt cccgcagaac 1500
gtgctgtctc tgttacgt 1518

<210> 179
<211> 299
<212> PRT
<213> Artificial Sequence

<220>
<223> HA1-1his (PR8)

<400> 179
Met Lys Phe Leu Val Asn Val Ala Leu Val Phe Met Val Val Tyr Ile
 1   5   10   15
Ser Tyr Ile Tyr Ala Ser His Asn Gly Lys Leu Cys Arg Leu Lys Gly
            20   25   30
Ile Ala Pro Leu Gln Leu Gly Lys Cys Asn Ile Ala Gly Trp Leu Leu
        35   40   45
Gly Asn Pro Glu Cys Asp Pro Leu Leu Pro Val Arg Ser Trp Ser Tyr
    50   55   60
Ile Val Glu Thr Pro Asn Ser Glu Asn Gly Ile Cys Tyr Pro Gly Asp
65   70   75   80
Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser
                85   90   95
Phe Glu Arg Phe Glu Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His
            100   105   110
Asn Thr Asn Gly Val Thr Ala Ala Cys Ser His Glu Gly Lys Ser Ser
        115   120   125
Phe Tyr Arg Asn Leu Leu Trp Leu Thr Glu Lys Glu Gly Ser Tyr Pro
    130   135   140

```
Lys Leu Lys Asn Ser Tyr Val Asn Lys Lys Gly Lys Glu Val Leu Val
145                 150                 155                 160
Leu Trp Gly Ile His His Pro Pro Asn Ser Lys Glu Gln Gln Asn Leu
                165                 170                 175
Tyr Gln Asn Glu Asn Ala Tyr Val Ser Val Val Thr Ser Asn Tyr Asn
            180                 185                 190
Arg Arg Phe Thr Pro Glu Ile Ala Glu Arg Pro Lys Val Arg Asp Gln
        195                 200                 205
Ala Gly Arg Met Asn Tyr Tyr Trp Thr Leu Leu Lys Pro Gly Asp Thr
    210                 215                 220
Ile Ile Phe Glu Ala Asn Gly Asn Leu Ile Ala Pro Met Tyr Ala Phe
225                 230                 235                 240
Ala Leu Ser Arg Gly Phe Gly Ser Gly Ile Ile Thr Ser Asn Ala Ser
                245                 250                 255
Met His Glu Cys Asn Thr Lys Cys Gln Thr Pro Leu Gly Ala Ile Asn
            260                 265                 270
Ser Ser Leu Pro Tyr Gln Asn Ile His Pro Val Thr Ile Gly Glu Cys
        275                 280                 285
Pro Lys Tyr Val Arg His His His His His His
    290                 295


<210> 180
<211> 301
<212> PRT
<213> Artificial Sequence

<220>
<223> HA1-1his (VN)

<400> 180
Met Lys Phe Leu Val Asn Val Ala Leu Val Phe Met Val Val Tyr Ile
 1               5                  10                  15
Ser Tyr Ile Tyr Ala Glu Lys Lys His Asn Gly Lys Leu Cys Asp Leu
            20                  25                  30
Asp Gly Val Lys Pro Leu Ile Leu Arg Asp Cys Ser Val Ala Gly Trp
        35                  40                  45
Leu Leu Gly Asn Pro Met Cys Asp Glu Phe Ile Asn Val Pro Glu Trp
    50                  55                  60
Ser Tyr Ile Val Glu Lys Ala Asn Pro Val Asn Asp Leu Cys Tyr Pro
65                  70                  75                  80
Gly Asp Phe Asn Asp Tyr Glu Glu Leu Lys His Leu Leu Ser Arg Ile
                85                  90                  95
Asn His Phe Glu Lys Ile Gln Ile Ile Pro Lys Ser Ser Trp Ser Ser
            100                 105                 110
His Glu Ala Ser Leu Gly Val Ser Ser Ala Cys Pro Tyr Gln Gly Lys
        115                 120                 125
Ser Ser Phe Phe Arg Asn Val Val Trp Leu Ile Lys Lys Asn Ser Thr
    130                 135                 140
Tyr Pro Thr Ile Lys Arg Ser Tyr Asn Asn Thr Asn Gln Glu Asp Leu
145                 150                 155                 160
Leu Val Leu Trp Gly Ile His His Pro Asn Asp Ala Ala Glu Gln Thr
                165                 170                 175
Lys Leu Tyr Gln Asn Pro Thr Thr Tyr Ile Ser Val Gly Thr Ser Thr
            180                 185                 190
Leu Asn Gln Arg Leu Val Pro Arg Ile Ala Thr Arg Ser Lys Val Asn
        195                 200                 205
Gly Gln Ser Gly Arg Met Glu Phe Phe Trp Thr Ile Leu Lys Pro Asn
    210                 215                 220
```

```
Asp Ala Ile Asn Phe Glu Ser Asn Gly Asn Phe Ile Ala Pro Glu Tyr
225                 230                 235                 240
Ala Tyr Lys Ile Val Lys Lys Gly Asp Ser Thr Ile Met Lys Ser Glu
                245                 250                 255
Leu Glu Tyr Gly Asn Cys Asn Thr Lys Cys Gln Thr Pro Met Gly Ala
            260                 265                 270
Ile Asn Ser Ser Met Pro Phe His Asn Ile His Pro Leu Thr Ile Gly
        275                 280                 285
Glu Cys Pro Lys Tyr Val Lys His His His His His His
    290                 295                 300


<210> 181
<211> 301
<212> PRT
<213> Artificial Sequence

<220>
<223> HA1-1his (IND)

<400> 181
Met Lys Phe Leu Val Asn Val Ala Leu Val Phe Met Val Val Tyr Ile
 1               5                  10                  15
Ser Tyr Ile Tyr Ala Glu Lys Thr His Asn Gly Lys Leu Cys Asp Leu
            20                  25                  30
Asp Gly Val Lys Pro Leu Ile Leu Arg Asp Cys Ser Val Ala Gly Trp
        35                  40                  45
Leu Leu Gly Asn Pro Met Cys Asp Glu Phe Ile Asn Val Pro Glu Trp
    50                  55                  60
Ser Tyr Ile Val Glu Lys Ala Asn Pro Thr Asn Asp Leu Cys Tyr Pro
65                  70                  75                  80
Gly Ser Phe Asn Asp Tyr Glu Glu Leu Lys His Leu Leu Ser Arg Ile
                85                  90                  95
Asn His Phe Glu Lys Ile Gln Ile Ile Pro Lys Ser Ser Trp Ser Asp
            100                 105                 110
His Glu Ala Ser Ser Gly Val Ser Ser Ala Cys Pro Tyr Leu Gly Ser
        115                 120                 125
Pro Ser Phe Phe Arg Asn Val Val Trp Leu Ile Lys Lys Asn Ser Thr
    130                 135                 140
Tyr Pro Thr Ile Lys Lys Ser Tyr Asn Asn Thr Asn Gln Glu Asp Leu
145                 150                 155                 160
Leu Val Leu Trp Gly Ile His His Pro Asn Asp Ala Ala Glu Gln Thr
                165                 170                 175
Arg Leu Tyr Gln Asn Pro Thr Thr Tyr Ile Ser Ile Gly Thr Ser Thr
            180                 185                 190
Leu Asn Gln Arg Leu Val Pro Lys Ile Ala Thr Arg Ser Lys Val Asn
        195                 200                 205
Gly Gln Ser Gly Arg Met Glu Phe Phe Trp Thr Ile Leu Lys Pro Asn
    210                 215                 220
Asp Ala Ile Asn Phe Glu Ser Asn Gly Asn Phe Ile Ala Pro Glu Tyr
225                 230                 235                 240
Ala Tyr Lys Ile Val Lys Lys Gly Asp Ser Ala Ile Met Lys Ser Glu
                245                 250                 255
Leu Glu Tyr Gly Asn Cys Asn Thr Lys Cys Gln Thr Pro Met Gly Ala
            260                 265                 270
Ile Asn Ser Ser Met Pro Phe His Asn Ile His Pro Leu Thr Ile Gly
        275                 280                 285
Glu Cys Pro Lys Tyr Val Lys His His His His His His
    290                 295                 300
```

<210> 182
<211> 299
<212> PRT
<213> Artificial Sequence

<220>
<223> HA1-1his (NC)

<400> 182
Met Lys Phe Leu Val Asn Val Ala Leu Val Phe Met Val Val Tyr Ile
1 5 10 15
Ser Tyr Ile Tyr Ala Ser His Asn Gly Lys Leu Cys Leu Leu Lys Gly
20 25 30
Ile Ala Pro Leu Gln Leu Gly Asn Cys Ser Val Ala Gly Trp Ile Leu
35 40 45
Gly Asn Pro Glu Cys Glu Leu Leu Ile Ser Lys Glu Ser Trp Ser Tyr
50 55 60
Ile Val Glu Thr Pro Asn Pro Glu Asn Gly Thr Cys Tyr Pro Gly Tyr
65 70 75 80
Phe Ala Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser
85 90 95
Phe Glu Arg Phe Glu Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His
100 105 110
Thr Val Thr Gly Val Ser Ala Ser Cys Ser His Asn Gly Lys Ser Ser
115 120 125
Phe Tyr Arg Asn Leu Leu Trp Leu Thr Gly Lys Asn Gly Leu Tyr Pro
130 135 140
Asn Leu Ser Lys Ser Tyr Val Asn Asn Lys Glu Lys Glu Val Leu Val
145 150 155 160
Leu Trp Gly Val His His Pro Pro Asn Ile Gly Asn Gln Arg Ala Leu
165 170 175
Tyr His Thr Glu Asn Ala Tyr Val Ser Val Val Ser Ser His Tyr Ser
180 185 190
Arg Arg Phe Thr Pro Glu Ile Ala Lys Arg Pro Lys Val Arg Asp Gln
195 200 205
Glu Gly Arg Ile Asn Tyr Tyr Trp Thr Leu Leu Glu Pro Gly Asp Thr
210 215 220
Ile Ile Phe Glu Ala Asn Gly Asn Leu Ile Ala Pro Trp Tyr Ala Phe
225 230 235 240
Ala Leu Ser Arg Gly Phe Gly Ser Gly Ile Ile Thr Ser Asn Ala Pro
245 250 255
Met Asp Glu Cys Asp Ala Lys Cys Gln Thr Pro Gln Gly Ala Ile Asn
260 265 270
Ser Ser Leu Pro Phe Gln Asn Val His Pro Val Thr Ile Gly Glu Cys
275 280 285
Pro Lys Tyr Val Arg His His His His His His
290 295

<210> 183
<211> 888
<212> DNA
<213> Artificial Sequence

<220>
<223> HA1-1 (MAL)

<400> 183
accactaccc caaccaaatc tcacttcgca aacctgaaag gcactgaaac ccgtggcaag 60
ctgtgtccga agtgtctgaa ctgcaccgat ctggacgtcg cactgggtcg tccgaaatgt 120
actggtaaca ttccgtccgc gcgtgtctcc atcctgcatg aagtgcgtcc agtgacctcc 180
ggctgttttc cgattatgca tgatcgtact aaaatccgtc agctgccgaa cctgctgcgt 240
ggttacgaac acattcgtct gtccacccat aacgttatca acgcggaaaa cgcgccgggc 300
ggtagctata aaatcggtac ctctggttct tgcccgaacg tgactaacgg taacggcttc 360
tttgcaacca tggcctgggc ggtcccgaaa aacgacaaca acaagaccgc gaccaattcc 420
ctgaccatcg aagtcccgta tatctgcacc gaaggtgaag atcaaatcac ggtttggggc 480
ttccactccg acaacgaggc acaaatggcg aaactgtacg gtgacagcaa accgcaaaaa 540
ttcactagct ccgctaacgg tgttaccacc cactacgttt cccagatcgg tggtttccca 600
aaccagaccg aagatggtgg tctgccgcag tccggtcgca tcgttgtaga ttatatggtg 660
cagaaaagcg gtaaaaccgg taccatcacc taccagcgtg gcatcctgct gccgcagaaa 720
gtttggtgcg cttccggtcg tagcaaagta atcaaaggtt ccctgccgct gatcggtgaa 780
gcagactgcc tgcacgagaa atacggcggt ctgaacaaaa gcaagccgta ctataccggc 840
gaacatgcga aagcaattgg taactgtcca atttgggtga aatagtag 888

<210> 184
<211> 2406
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2.HA1-1 (MAL)

<400> 184
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcagaa agcgtatgat 540
gtgaaagata cagcagtaac aacgaaagct tatgccaata atggtactac actggatgta 600
tcgggtcttg atgatgcagc tattaaagcg gctacgggtg gtacgaatgg tacggcttct 660
gtaaccggtg gtgcggttaa atttgacgca gataataaca agtactttgt tactattggt 720
ggctttactg gtgctgatgc cgccaaaaat ggcgattatg aagttaacgt tgctactgac 780
ggtacagtaa cccttgcggc tggcgcaact aaaaccacaa tgcctgctgg tgcgacaact 840
aaaacagaag tacaggagtt aaaagataca ccggcagttg tttcagcaga tgctaaaaat 900
gccttaattg ctggcggcgt tgacgctacc gatgctaatg gcgctgagtt ggtcaaaatg 960
tcttataccg ataaaaatgg taagacaatt gaaggcggtt atgcgcttaa agctggcgat 1020
aagtattacg ccgcagatta cgatgaagcg acaggagcaa ttaaagctaa aactacaagt 1080
tatactgctg ctgacggcac taccaaaaca gcggctaacc aactgggtgg cgtagacggt 1140
aaaaccgaag tcgttactat cgacggtaaa acctacaatg ccagcaaagc cgctggtcat 1200
gatttcaaag cacaaccaga gctggcggaa gcagccgcta aaaccaccga aaacccgctg 1260
cagaaaattg atgccgcgct ggcgcaggtg gatgcgctgc gctctgatct gggtgcggta 1320
caaaaccgtt tcaactctgc tatcaccaac ctgggcaata ccgtaaacaa tctgtctgaa 1380
gcgcgtagcc gtatcgaaga ttccgactac gcgaccgaag tttccaacat gtctcgcgcg 1440
cagattctgc agcaggccgg tacttccgtt ctggcgcagg ctaaccaggt cccgcagaac 1500
gtgctgtctc tgttagcgac cactacccca accaaatctc acttcgcaaa cctgaaaggc 1560
actgaaaccc gtggcaagct gtgtccgaag tgtctgaact gcaccgatct ggacgtcgca 1620
ctgggtcgtc cgaaatgtac tggtaacatt ccgtccgcgc gtgtctccat cctgcatgaa 1680
gtgcgtccag tgacctccgg ctgttttccg attatgcatg atcgtactaa aatccgtcag 1740
ctgccgaacc tgctgcgtgg ttacgaacac attcgtctgt ccacccataa cgttatcaac 1800
gcggaaaacg cgccgggcgg tagctataaa atcggtacct ctggttcttg cccgaacgtg 1860
actaacggta acggcttctt tgcaaccatg gcctgggcgg tcccgaaaaa cgacaacaac 1920
aagaccgcga ccaattccct gaccatcgaa gtcccgtata tctgcaccga aggtgaagat 1980

caaatcacgg tttggggctt ccactccgac aacgaggcac aaatggcgaa actgtacggt 2040
gacagcaaac cgcaaaaatt cactagctcc gctaacggtg ttaccaccca ctacgtttcc 2100
cagatcggtg gtttcccaaa ccagaccgaa gatggtggtc tgccgcagtc cggtcgcatc 2160
gttgtagatt atatggtgca gaaaagcggt aaaaccggta ccatcaccta ccagcgtggc 2220
atcctgctgc cgcagaaagt ttggtgcgct tccggtcgta gcaaagtaat caaaggttcc 2280
ctgccgctga tcggtgaagc agactgcctg cacgagaaat acggcggtct gaacaaaagc 2340
aagccgtact ataccggcga acatgcgaaa gcaattggta actgtccaat ttgggtgaaa 2400
tagtag 2406

<210> 185
<211> 729
<212> DNA
<213> Artificial Sequence

<220>
<223> HA1-2 (MAL)

<400> 185
aaaggcactg aaacccgtgg caagctgtgt ccgaagtgtc tgaactgcac cgatctggac 60
gtcgcactgg gtcgtccgaa atgtactggt aacattccgt ccgcgcgtgt ctccatcctg 120
catgaagtgc gtccagtgac ctccggctgt tttccgatta tgcatgatcg tactaaaatc 180
cgtcagctgc cgaacctgct gcgtggttac gaacacattc gtctgtccac ccataacgtt 240
atcaacgcgg aaaacgcgcc gggcggtagc tataaaatcg gtacctctgg ttcttgcccg 300
aacgtgacta acggtaacgg cttctttgca accatggcct gggcggtccc gaaaaacgac 360
aacaacaaga ccgcgaccaa ttccctgacc atcgaagtcc cgtatatctg caccgaaggt 420
gaagatcaaa tcacggtttg gggcttccac tccgacaacg aggcacaaat ggcgaaactg 480
tacggtgaca gcaaaccgca aaaattcact agctccgcta acggtgttac cacccactac 540
gtttcccaga tcggtggttt cccaaaccag accgaagatg gtggtctgcc gcagtccggt 600
cgcatcgttg tagattatat ggtgcagaaa agcggtaaaa ccggtaccat cacctaccag 660
cgtggcatcc tgctgccgca gaaagtttgg tgcgcttccg gtcgtagcaa agtaatcaaa 720
ggttgatag 729

<210> 186
<211> 2247
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2.HA1-2 (MAL)

<400> 186
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcagaa agcgtatgat 540
gtgaaagata cagcagtaac aacgaaagct tatgccaata atggtactac actggatgta 600
tcgggtcttg atgatgcagc tattaaagcg gctacgggtg gtacgaatgg tacggcttct 660
gtaaccggtg gtgcggttaa atttgacgca gataataaca agtactttgt tactattggt 720
ggctttactg gtgctgatgc cgccaaaaat ggcgattatg aagttaacgt tgctactgac 780
ggtacagtaa cccttgcggc tggcgcaact aaaaccacaa tgcctgctgg tgcgacaact 840
aaaacagaag tacaggagtt aaaagataca ccggcagttg tttcagcaga tgctaaaaat 900
gccttaattg ctggcggcgt tgacgctacc gatgctaatg gcgctgagtt ggtcaaaatg 960
tcttataccg ataaaaatgg taagacaatt gaaggcggtt atgcgcttaa agctggcgat 1020
aagtattacg ccgcagatta cgatgaagcg acaggagcaa ttaaagctaa aactacaagt 1080

tatactgctg ctgacggcac taccaaaaca gcggctaacc aactgggtgg cgtagacggt 1140
aaaaccgaag tcgttactat cgacggtaaa acctacaatg ccagcaaagc cgctggtcat 1200
gatttcaaag cacaaccaga gctggcggaa gcagccgcta aaaccaccga aaacccgctg 1260
cagaaaattg atgccgcgct ggcgcaggtg gatgcgctgc gctctgatct gggtgcggta 1320
caaaaccgtt tcaactctgc tatcaccaac ctgggcaata ccgtaaacaa tctgtctgaa 1380
gcgcgtagcc gtatcgaaga ttccgactac gcgaccgaag tttccaacat gtctcgcgcg 1440
cagattctgc agcaggccgg tacttccgtt ctggcgcagg ctaaccaggt cccgcagaac 1500
gtgctgtctc tgttagcgaa aggcactgaa acccgtggca agctgtgtcc gaagtgtctg 1560
aactgcaccg atctggacgt cgcactgggt cgtccgaaat gtactggtaa cattccgtcc 1620
gcgcgtgtct ccatcctgca tgaagtgcgt ccagtgacct ccggctgttt tccgattatg 1680
catgatcgta ctaaaatccg tcagctgccg aacctgctgc gtggttacga acacattcgt 1740
ctgtccaccc ataacgttat caacgcggaa aacgcgccgg gcggtagcta taaaatcggt 1800
acctctggtt cttgcccgaa cgtgactaac ggtaacggct tctttgcaac catggcctgg 1860
gcggtcccga aaaacgacaa caacaagacc gcgaccaatt ccctgaccat cgaagtcccg 1920
tatatctgca ccgaaggtga agatcaaatc acggtttggg gcttccactc cgacaacgag 1980
gcacaaatgg cgaaactgta cggtgacagc aaaccgcaaa aattcactag ctccgctaac 2040
ggtgttacca cccactacgt ttcccagatc ggtggtttcc caaaccagac cgaagatggt 2100
ggtctgccgc agtccggtcg catcgttgta gattatatgg tgcagaaaag cggtaaaacc 2160
ggtaccatca cctaccagcg tggcatcctg ctgccgcaga aagtttggtg cgcttccggt 2220
cgtagcaaag taatcaaagg ttgatag 2247

<210> 187
<211> 725
<212> DNA
<213> Artificial Sequence

<220>
<223> HA1-2 (SH)

<400> 187
aaaggcactc gtacccgcgg taagctgtgc ccggactgcc tgaactgtac cgatctggat 60
gttgcactgg gtcgtccgat gtgcgttggt accacccccgt ctgcgaaagc cagcatcctg 120
cacgaagttc gcccggttac ttccggttgt ttcccgatta tgcatgatcg taccaaaatt 180
cgtcagctgc caaacctgct gcgtggctat gaaaacattc gtctgtccac tcaaaacgta 240
atcgatgcag aaaaagcgct gggtggcccg tatcgtctgg gtaccagcgg ctcctgcccg 300
aacgcgacga gcaaaagcgg cttcttcgcc accatggctt gggccgttcc gaaagacaac 360
aacaaaaacg ctacgaaccc gctgaccgtc gaagtcccgt acatctgcac cgaaggcgaa 420
gatcagatca ctgtgtgggg cttccacagc gatgataaga cccagatgaa aaatctgtac 480
ggtgactcca acccgcagaa attcacctct tctgctaacg gtgtaacgac ccactacgtt 540
tctcagatcg gtggtttccc ggaccagacg gaagatggcg gtctgcctca gtccggccgc 600
atcgtagttg attacatggt ccagaaaccg ggtaagactg gtaccattgt ttaccagcgt 660
ggtgtactgc tgccgcagaa ggtctggtgt gcttccggcc gttccaaggt cattaagggc 720
tgata 725

<210> 188
<211> 2244
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2.HA1-2 (SH)

<400> 188
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360

aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcagaa agcgtatgat 540
gtgaaagata cagcagtaac aacgaaagct tatgccaata atggtactac actggatgta 600
tcgggtcttg atgatgcagc tattaaagcg gctacgggtg gtacgaatgg tacggcttct 660
gtaaccggtg gtgcggttaa atttgacgca gataataaca agtactttgt tactattggt 720
ggctttactg gtgctgatgc cgccaaaaat ggcgattatg aagttaacgt tgctactgac 780
ggtacagtaa cccttgcggc tggcgcaact aaaaccacaa tgcctgctgg tgcgacaact 840
aaaacagaag tacaggagtt aaaagataca ccggcagttg tttcagcaga tgctaaaaat 900
gccttaattg ctggcggcgt tgacgctacc gatgctaatg gcgctgagtt ggtcaaaatg 960
tcttataccg ataaaaatgg taagacaatt gaaggcggtt atgcgcttaa agctggcgat 1020
aagtattacg ccgcagatta cgatgaagcg acaggagcaa ttaaagctaa aactacaagt 1080
tatactgctg ctgacggcac taccaaaaca gcggctaacc aactgggtgg cgtagacggt 1140
aaaaccgaag tcgttactat cgacggtaaa acctacaatg ccagcaaagc cgctggtcat 1200
gatttcaaag cacaaccaga gctggcggaa gcagccgcta aaaccaccga aaacccgctg 1260
cagaaaattg atgccgcgct ggcgcaggtg gatgcgctgc gctctgatct gggtgcggta 1320
caaaaccgtt tcaactctgc tatcaccaac ctgggcaata ccgtaaacaa tctgtctgaa 1380
gcgcgtagcc gtatcgaaga ttccgactac gcgaccgaag tttccaacat gtctcgcgcg 1440
cagattctgc agcaggccgg tacttccgtt ctggcgcagg ctaaccaggt cccgcagaac 1500
gtgctgtctc tgttagcgaa aggcactcgt acccgcggta agctgtgccc ggactgcctg 1560
aactgtaccg atctggatgt tgcactgggt cgtccgatgt gcgttggtac caccccgtct 1620
gcgaaagcca gcatcctgca cgaagttcgc ccggttactt ccggttgttt cccgattatg 1680
catgatcgta ccaaaattcg tcagctgcca aacctgctgc gtggctatga aaacattcgt 1740
ctgtccactc aaaacgtaat cgatgcagaa aaagcgctgg gtggcccgta tcgtctgggt 1800
accagcggct cctgcccgaa cgcgacgagc aaaagcggct tcttcgccac catggcttgg 1860
gccgttccga aagacaacaa caaaaacgct acgaacccgc tgaccgtcga agtcccgtac 1920
atctgcaccg aaggcgaaga tcagatcact gtgtggggct tccacagcga tgataagacc 1980
cagatgaaaa atctgtacgg tgactccaac ccgcagaaat tcacctcttc tgctaacggt 2040
gtaacgaccc actacgtttc tcagatcggt ggtttcccgg accagacgga agatggcggt 2100
ctgcctcagt ccggccgcat cgtagttgat tacatggtcc agaaaccggg taagactggt 2160
accattgttt accagcgtgg tgtactgctg ccgcagaagg tctggtgtgc ttccggccgt 2220
tccaaggtca ttaagggctg atag 2244

<210> 189
<211> 726
<212> DNA
<213> Artificial Sequence

<220>
<223> HA1-2 (Lee)

<400> 189
aaaggcactc agacccgtgg caagctgtgt ccgaactgtt tcaactgcac cgatctggac 60
gttgcactgg gtcgtccgaa atgcatgggt aacatcccgt ctgcgaaggt aagcatcctg 120
cacgaagtta aaccggtaac cagcggctgt ttcccgatca tgcacgacaa aactaaaatt 180
cgtcagctgc cgaacctgct gcgtggttat gagaacattc gtctgtctac ctctaatgtt 240
atcaacgcgg agactgcacc aggtggccca tacaaagtag gtaccagcgg ttcctgtccg 300
aacgttgcga atcgtaacgg cttcttcaac actatggcgt gggttatccc gaaagataac 360
aataaaactg caattaaccc ggtaactgta gaagttccgt acatctgctc cgaaggcgag 420
gaccagatta cggtatgggg ctttcacagc gacgataaaa cccagatgga gcgtctgtac 480
ggtgactcta acccgcagaa attcacctcc tccgcgaacg gcgttaccac ccactatgtt 540
tctcagatcg gcggtttccc gaatcagacc gaagacgaag gcctgaagca gtccggccgt 600
attgttgtag actacatggt tcagaagccg ggcaaaactg gtaccattgt ataccagcgc 660
ggcatcctgc tgccgcagaa agtttggtgc gcttccggtc gtagcaaagt aatcaaaggt 720
tgatag 726

<210> 190
<211> 2244
<212> DNA

<213> Artificial Sequence

<220>
<223> STF2.HA1-2 (Lee)

<400> 190
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcagaa agcgtatgat 540
gtgaaagata cagcagtaac aacgaaagct tatgccaata atggtactac actggatgta 600
tcgggtcttg atgatgcagc tattaaagcg gctacgggtg gtacgaatgg tacggcttct 660
gtaaccggtg gtgcggttaa atttgacgca gataataaca agtactttgt tactattggt 720
ggctttactg gtgctgatgc cgccaaaaat ggcgattatg aagttaacgt tgctactgac 780
ggtacagtaa cccttgcggc tggcgcaact aaaaccacaa tgcctgctgg tgcgacaact 840
aaaacagaag tacaggagtt aaaagataca ccggcagttg tttcagcaga tgctaaaaat 900
gccttaattg ctggcggcgt tgacgctacc gatgctaatg gcgctgagtt ggtcaaaatg 960
tcttataccg ataaaaatgg taagacaatt gaaggcggtt atgcgcttaa agctggcgat 1020
aagtattacg ccgcagatta cgatgaagcg acaggagcaa ttaaagctaa aactacaagt 1080
tatactgctg ctgacggcac taccaaaaca gcggctaacc aactgggtgg cgtagacggt 1140
aaaaccgaag tcgttactat cgacggtaaa acctacaatg ccagcaaagc cgctggtcat 1200
gatttcaaag cacaaccaga gctggcggaa gcagccgcta aaaccaccga aaacccgctg 1260
cagaaaattg atgccgcgct ggcgcaggtg gatgcgctgc gctctgatct gggtgcggta 1320
caaaaccgtt tcaactctgc tatcaccaac ctgggcaata ccgtaaacaa tctgtctgaa 1380
gcgcgtagcc gtatcgaaga ttccgactac gcgaccgaag tttccaacat gtctcgcgcg 1440
cagattctgc agcaggccgg tacttccgtt ctggcgcagg ctaaccaggt cccgcagaac 1500
gtgctgtctc tgttagcgaa aggcactcag acccgtggca agctgtgtcc gaactgtttc 1560
aactgcaccg atctggacgt tgcactgggt cgtccgaaat gcatgggtaa catcccgtct 1620
gcgaaggtaa gcatcctgca cgaagttaaa ccggtaacca gcggctgttt cccgatcatg 1680
cacgacaaaa ctaaaattcg tcagctgccg aacctgctgc gtggttatga gaacattcgt 1740
ctgtctacct ctaatgttat caacgcggag actgcaccag gtggcccata caaagtaggt 1800
accagcggtt cctgtccgaa cgttgcgaat cgtaacggct tcttcaacac tatggcgtgg 1860
gttatcccga aagataacaa taaaactgca attaacccgg taactgtaga agttccgtac 1920
atctgctccg aaggcgagga ccagattacg gtatggggct ttcacagcga cgataaaacc 1980
cagatggagc gtctgtacgg tgactctaac ccgcagaaat tcacctcctc cgcgaacggc 2040
gttaccaccc actatgtttc tcagatcggc ggtttcccga atcagaccga agacgaaggc 2100
ctgaagcagt ccggccgtat tgttgtagac tacatggttc agaagccggg caaaactggt 2160
accattgtat accagcgcgg catcctgctg ccgcagaaag tttggtgcgc ttccggtcgt 2220
agcaaagtaa tcaaaggttg atag 2244

<210> 191
<211> 729
<212> DNA
<213> Artificial Sequence

<220>
<223> HA1-2 (Ohio)

<400> 191
aaaggcacta aaacccgtgg caagctgtgt ccgaagtgtc tgaactgcac cgatctggac 60
gtcgcactgg gtcgtccgaa atgtactggt aacattccgt ccgcggaagt ctccatcctg 120
catgaagtgc gtccagtgac ctccggctgt tttccgatta tgcatgatcg tactaaaatc 180
cgtcagctgc cgaacctgct gcgtggttac gaacacattc gtctgtccac ccataacgtt 240
atcaacgcgg aaaaggcgcc gggcggtccc tataaaatcg gtacctctgg ttcttgcccg 300

aacgtgacta acggtaacgg cttctttgca accatggcct gggcggtccc gaaaaacgac 360
aacaacaaga ccgcgaccaa ttccctgacc atcgaagtcc cgtatatctg caccgaaggt 420
gaagatcaaa tcacgatttg gggcttccac tccgacagcg agacacaaat ggcgaaactg 480
tacggtgaca gcaaaccgca aaaattcact agctccgcta acggtgttac cacccactac 540
gtttcccaga tcggtggttt cccaaaccag accgaagatg gtggtctgcc gcagtccggt 600
cgcatcgttg tagattatat ggtgcagaaa agcggtaaaa ccggtaccat cacctaccag 660
cgtggcatcc tgctgccgca gaaagtttgg tgcgcttccg gtcgtagcaa agtaatcaaa 720
ggttgatag 729

<210> 192
<211> 2247
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2.HA1-2 (Ohio)

<400> 192
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcagaa agcgtatgat 540
gtgaaagata cagcagtaac aacgaaagct tatgccaata atggtactac actggatgta 600
tcgggtcttg atgatgcagc tattaaagcg gctacgggtg gtacgaatgg tacggcttct 660
gtaaccggtg gtgcggttaa atttgacgca gataataaca agtactttgt tactattggt 720
ggctttactg gtgctgatgc cgccaaaaat ggcgattatg aagttaacgt tgctactgac 780
ggtacagtaa cccttgcggc tggcgcaact aaaaccacaa tgcctgctgg tgcgacaact 840
aaaacagaag tacaggagtt aaaagataca ccggcagttg tttcagcaga tgctaaaaat 900
gccttaattg ctggcggcgt tgacgctacc gatgctaatg gcgctgagtt ggtcaaaatg 960
tcttataccg ataaaaatgg taagacaatt gaaggcggtt atgcgcttaa agctggcgat 1020
aagtattacg ccgcagatta cgatgaagcg acaggagcaa ttaaagctaa aactacaagt 1080
tatactgctg ctgacggcac taccaaaaca gcggctaacc aactgggtgg cgtagacggt 1140
aaaaccgaag tcgttactat cgacggtaaa acctacaatg ccagcaaagc cgctggtcat 1200
gatttcaaag cacaaccaga gctggcggaa gcagccgcta aaaccaccga aaacccgctg 1260
cagaaaattg atgccgcgct ggcgcaggtg gatgcgctgc gctctgatct gggtgcggta 1320
caaaaccgtt tcaactctgc tatcaccaac ctgggcaata ccgtaaacaa tctgtctgaa 1380
gcgcgtagcc gtatcgaaga ttccgactac gcgaccgaag tttccaacat gtctcgcgcg 1440
cagattctgc agcaggccgg tacttccgtt ctggcgcagg ctaaccaggt cccgcagaac 1500
gtgctgtctc tgttagcgaa aggcactaaa acccgtggca agctgtgtcc gaagtgtctg 1560
aactgcaccg atctggacgt cgcactgggt cgtccgaaat gtactggtaa cattccgtcc 1620
gcggaagtct ccatcctgca tgaagtgcgt ccagtgacct ccggctgttt tccgattatg 1680
catgatcgta ctaaaatccg tcagctgccg aacctgctgc gtggttacga acacattcgt 1740
ctgtccaccc ataacgttat caacgcggaa aaggcgccgg gcggtcccta taaaatcggt 1800
acctctggtt cttgcccgaa cgtgactaac ggtaacggct tctttgcaac catggcctgg 1860
gcggtcccga aaaacgacaa caacaagacc gcgaccaatt ccctgaccat cgaagtcccg 1920
tatatctgca ccgaaggtga agatcaaatc acgatttggg gcttccactc cgacagcgag 1980
acacaaatgg cgaaactgta cggtgacagc aaaccgcaaa aattcactag ctccgctaac 2040
ggtgttacca cccactacgt ttcccagatc ggtggtttcc caaaccagac cgaagatggt 2100
ggtctgccgc agtccggtcg catcgttgta gattatatgg tgcagaaaag cggtaaaacc 2160
ggtaccatca cctaccagcg tggcatcctg ctgccgcaga aagtttggtg cgcttccggt 2220
cgtagcaaag taatcaaagg ttgatag 2247

<210> 193
<211> 21

<212> DNA
<213> Artificial Sequence

<220>
<223> FOR PRIMER

<400> 193
cggataacaa ttcccctcta g 21

<210> 194
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> REV PRIMER

<400> 194
cgaagtgaga tttggttgga gtagtggtcg ctaacagaga cagcacgttc 50

<210> 195
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> FOR PRIMER

<400> 195
cccgcagaac gtgctgtctc tgttagcgac cactacccca accaaatctc 50

<210> 196
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> REV PRIMER

<400> 196
ctatttcacc caaattggac 20

<210> 197
<211> 51
<212> DNA
<213> Artificial Sequence

<220>
<223> REV PRIMER

<400> 197
cacagcttgc cacgggtttc agtgcctttc gctaacagag acagcacgtt c 51

<210> 198
<211> 49
<212> DNA
<213> Artificial Sequence

<220>

<223> FOR PRIMER

<400> 198
cccgcagaac gtgctgtctc tgttagcgaa aggcactgaa acccgtggc 49

<210> 199
<211> 44
<212> DNA
<213> Artificial Sequence

<220>
<223> REV PRIMER

<400> 199
gactagacgc tcagctatca acctttgatt actttgctac gacc 44

<210> 200
<211> 51
<212> DNA
<213> Artificial Sequence

<220>
<223> REV PRIMER

<400> 200
cacagcttac cgcgggtacg agtgcctttc gctaacagag acagcacgtt c 51

<210> 201
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> FOR PRIMER

<400> 201
cgggatccaa aggcactcgt acccgcggta ag 32

<210> 202
<211> 44
<212> DNA
<213> Artificial Sequence

<220>
<223> REV PRIMER

<400> 202
gactagacgc tcagctatca gcccttaatg accttggaac ggcc 44

<210> 203
<211> 52
<212> DNA
<213> Artificial Sequence

<220>
<223> REV PRIMER

<400> 203
gcatagtttt cctctggtct gtgttccttt cgctaacaga gacagcacgt tc 52

<210> 204
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> FOR PRIMER

<400> 204
aaaggaacac agaccagagg 20

<210> 205
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> REV PRIMER

<400> 205
ctactaccct tttattacct tgctcc 26

<210> 206
<211> 51
<212> DNA
<213> Artificial Sequence

<220>
<223> REV PRIMER

<400> 206
cacagcttgc cacgggtttt agtgcctttc gctaacagag acagcacgtt c 51

<210> 207
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> FOR PRIMER

<400> 207
aaaggcacta aaacccgtgg c 21

<210> 208
<211> 44
<212> DNA
<213> Artificial Sequence

<220>
<223> REV PRIMER

<400> 208
gactagacgc tcagctatca acctttgatt actttgctac gacc 44

<210> 209
<211> 800
<212> PRT

<213> Artificial Sequence

<220>
<223> STF2.HA1-1 (MAL)

<400> 209

```
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
 1               5                  10                  15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
            20                  25                  30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
        35                  40                  45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
    50                  55                  60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65                  70                  75                  80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
                85                  90                  95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
            100                 105                 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
        115                 120                 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
    130                 135                 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145                 150                 155                 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val Gln
                165                 170                 175
Lys Ala Tyr Asp Val Lys Asp Thr Ala Val Thr Thr Lys Ala Tyr Ala
            180                 185                 190
Asn Asn Gly Thr Thr Leu Asp Val Ser Gly Leu Asp Asp Ala Ala Ile
        195                 200                 205
Lys Ala Ala Thr Gly Gly Thr Asn Gly Thr Ala Ser Val Thr Gly Gly
    210                 215                 220
Ala Val Lys Phe Asp Ala Asp Asn Asn Lys Tyr Phe Val Thr Ile Gly
225                 230                 235                 240
Gly Phe Thr Gly Ala Asp Ala Ala Lys Asn Gly Asp Tyr Glu Val Asn
                245                 250                 255
Val Ala Thr Asp Gly Thr Val Thr Leu Ala Ala Gly Ala Thr Lys Thr
            260                 265                 270
Thr Met Pro Ala Gly Ala Thr Thr Lys Thr Glu Val Gln Glu Leu Lys
        275                 280                 285
Asp Thr Pro Ala Val Val Ser Ala Asp Ala Lys Asn Ala Leu Ile Ala
    290                 295                 300
Gly Gly Val Asp Ala Thr Asp Ala Asn Gly Ala Glu Leu Val Lys Met
305                 310                 315                 320
Ser Tyr Thr Asp Lys Asn Gly Lys Thr Ile Glu Gly Gly Tyr Ala Leu
                325                 330                 335
Lys Ala Gly Asp Lys Tyr Tyr Ala Ala Asp Tyr Asp Glu Ala Thr Gly
            340                 345                 350
Ala Ile Lys Ala Lys Thr Thr Ser Tyr Thr Ala Ala Asp Gly Thr Thr
        355                 360                 365
Lys Thr Ala Ala Asn Gln Leu Gly Gly Val Asp Gly Lys Thr Glu Val
    370                 375                 380
Val Thr Ile Asp Gly Lys Thr Tyr Asn Ala Ser Lys Ala Ala Gly His
385                 390                 395                 400
Asp Phe Lys Ala Gln Pro Glu Leu Ala Glu Ala Ala Ala Lys Thr Thr
                405                 410                 415
Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala
```

```
            420                 425                 430
Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile
        435                 440                 445
Thr Asn Leu Gly Asn Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg
    450                 455                 460
Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala
465                 470                 475                 480
Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln
                485                 490                 495
Val Pro Gln Asn Val Leu Ser Leu Leu Ala Thr Thr Thr Pro Thr Lys
            500                 505                 510
Ser His Phe Ala Asn Leu Lys Gly Thr Glu Thr Arg Gly Lys Leu Cys
        515                 520                 525
Pro Lys Cys Leu Asn Cys Thr Asp Leu Asp Val Ala Leu Gly Arg Pro
    530                 535                 540
Lys Cys Thr Gly Asn Ile Pro Ser Ala Arg Val Ser Ile Leu His Glu
545                 550                 555                 560
Val Arg Pro Val Thr Ser Gly Cys Phe Pro Ile Met His Asp Arg Thr
                565                 570                 575
Lys Ile Arg Gln Leu Pro Asn Leu Leu Arg Gly Tyr Glu His Ile Arg
            580                 585                 590
Leu Ser Thr His Asn Val Ile Asn Ala Glu Asn Ala Pro Gly Gly Ser
        595                 600                 605
Tyr Lys Ile Gly Thr Ser Gly Ser Cys Pro Asn Val Thr Asn Gly Asn
    610                 615                 620
Gly Phe Phe Ala Thr Met Ala Trp Ala Val Pro Lys Asn Asp Asn Asn
625                 630                 635                 640
Lys Thr Ala Thr Asn Ser Leu Thr Ile Glu Val Pro Tyr Ile Cys Thr
                645                 650                 655
Glu Gly Glu Asp Gln Ile Thr Val Trp Gly Phe His Ser Asp Asn Glu
            660                 665                 670
Ala Gln Met Ala Lys Leu Tyr Gly Asp Ser Lys Pro Gln Lys Phe Thr
        675                 680                 685
Ser Ser Ala Asn Gly Val Thr Thr His Tyr Val Ser Gln Ile Gly Gly
    690                 695                 700
Phe Pro Asn Gln Thr Glu Asp Gly Gly Leu Pro Gln Ser Gly Arg Ile
705                 710                 715                 720
Val Val Asp Tyr Met Val Gln Lys Ser Gly Lys Thr Gly Thr Ile Thr
                725                 730                 735
Tyr Gln Arg Gly Ile Leu Leu Pro Gln Lys Val Trp Cys Ala Ser Gly
            740                 745                 750
Arg Ser Lys Val Ile Lys Gly Ser Leu Pro Leu Ile Gly Glu Ala Asp
        755                 760                 765
Cys Leu His Glu Lys Tyr Gly Gly Leu Asn Lys Ser Lys Pro Tyr Tyr
    770                 775                 780
Thr Gly Glu His Ala Lys Ala Ile Gly Asn Cys Pro Ile Trp Val Lys
785                 790                 795                 800


<210> 210
<211> 747
<212> PRT
<213> Artificial Sequence

<220>
<223> STF2.HA1-2 (MAL)

<400> 210
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
```

```
 1               5                  10                  15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
            20                  25                  30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
        35                  40                  45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
    50                  55                  60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65                  70                  75                  80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
                85                  90                  95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
            100                 105                 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
        115                 120                 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
    130                 135                 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145                 150                 155                 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val Gln
                165                 170                 175
Lys Ala Tyr Asp Val Lys Asp Thr Ala Val Thr Thr Lys Ala Tyr Ala
            180                 185                 190
Asn Asn Gly Thr Thr Leu Asp Val Ser Gly Leu Asp Asp Ala Ala Ile
        195                 200                 205
Lys Ala Ala Thr Gly Gly Thr Asn Gly Thr Ala Ser Val Thr Gly Gly
    210                 215                 220
Ala Val Lys Phe Asp Ala Asp Asn Asn Lys Tyr Phe Val Thr Ile Gly
225                 230                 235                 240
Gly Phe Thr Gly Ala Asp Ala Ala Lys Asn Gly Asp Tyr Glu Val Asn
                245                 250                 255
Val Ala Thr Asp Gly Thr Val Thr Leu Ala Ala Gly Ala Thr Lys Thr
            260                 265                 270
Thr Met Pro Ala Gly Ala Thr Thr Lys Thr Glu Val Gln Glu Leu Lys
        275                 280                 285
Asp Thr Pro Ala Val Val Ser Ala Asp Ala Lys Asn Ala Leu Ile Ala
    290                 295                 300
Gly Gly Val Asp Ala Thr Asp Ala Asn Gly Ala Glu Leu Val Lys Met
305                 310                 315                 320
Ser Tyr Thr Asp Lys Asn Gly Lys Thr Ile Glu Gly Gly Tyr Ala Leu
                325                 330                 335
Lys Ala Gly Asp Lys Tyr Tyr Ala Ala Asp Tyr Asp Glu Ala Thr Gly
            340                 345                 350
Ala Ile Lys Ala Lys Thr Thr Ser Tyr Thr Ala Ala Asp Gly Thr Thr
        355                 360                 365
Lys Thr Ala Ala Asn Gln Leu Gly Gly Val Asp Gly Lys Thr Glu Val
    370                 375                 380
Val Thr Ile Asp Gly Lys Thr Tyr Asn Ala Ser Lys Ala Ala Gly His
385                 390                 395                 400
Asp Phe Lys Ala Gln Pro Glu Leu Ala Glu Ala Ala Ala Lys Thr Thr
                405                 410                 415
Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala
            420                 425                 430
Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile
        435                 440                 445
Thr Asn Leu Gly Asn Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg
    450                 455                 460
Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala
465                 470                 475                 480
```

```
Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln
                485                 490                 495
Val Pro Gln Asn Val Leu Ser Leu Leu Ala Lys Gly Thr Glu Thr Arg
            500                 505                 510
Gly Lys Leu Cys Pro Lys Cys Leu Asn Cys Thr Asp Leu Asp Val Ala
        515                 520                 525
Leu Gly Arg Pro Lys Cys Thr Gly Asn Ile Pro Ser Ala Arg Val Ser
    530                 535                 540
Ile Leu His Glu Val Arg Pro Val Thr Ser Gly Cys Phe Pro Ile Met
545                 550                 555                 560
His Asp Arg Thr Lys Ile Arg Gln Leu Pro Asn Leu Leu Arg Gly Tyr
                565                 570                 575
Glu His Ile Arg Leu Ser Thr His Asn Val Ile Asn Ala Glu Asn Ala
            580                 585                 590
Pro Gly Gly Ser Tyr Lys Ile Gly Thr Ser Gly Ser Cys Pro Asn Val
        595                 600                 605
Thr Asn Gly Asn Gly Phe Phe Ala Thr Met Ala Trp Ala Val Pro Lys
    610                 615                 620
Asn Asp Asn Asn Lys Thr Ala Thr Asn Ser Leu Thr Ile Glu Val Pro
625                 630                 635                 640
Tyr Ile Cys Thr Glu Gly Glu Asp Gln Ile Thr Val Trp Gly Phe His
                645                 650                 655
Ser Asp Asn Glu Ala Gln Met Ala Lys Leu Tyr Gly Asp Ser Lys Pro
            660                 665                 670
Gln Lys Phe Thr Ser Ser Ala Asn Gly Val Thr Thr His Tyr Val Ser
        675                 680                 685
Gln Ile Gly Gly Phe Pro Asn Gln Thr Glu Asp Gly Gly Leu Pro Gln
    690                 695                 700
Ser Gly Arg Ile Val Val Asp Tyr Met Val Gln Lys Ser Gly Lys Thr
705                 710                 715                 720
Gly Thr Ile Thr Tyr Gln Arg Gly Ile Leu Leu Pro Gln Lys Val Trp
                725                 730                 735
Cys Ala Ser Gly Arg Ser Lys Val Ile Lys Gly
            740                 745


<210> 211
<211> 746
<212> PRT
<213> Artificial Sequence

<220>
<223> STF2.HA1-2 (SH)

<400> 211
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
 1               5                  10                  15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
            20                  25                  30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
        35                  40                  45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
    50                  55                  60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65                  70                  75                  80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
                85                  90                  95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
            100                 105                 110
```

```
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
        115                 120                 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
    130                 135                 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145                 150                 155                 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val Gln
                165                 170                 175
Lys Ala Tyr Asp Val Lys Asp Thr Ala Val Thr Thr Lys Ala Tyr Ala
            180                 185                 190
Asn Asn Gly Thr Thr Leu Asp Val Ser Gly Leu Asp Asp Ala Ala Ile
        195                 200                 205
Lys Ala Ala Thr Gly Gly Thr Asn Gly Thr Ala Ser Val Thr Gly Gly
    210                 215                 220
Ala Val Lys Phe Asp Ala Asp Asn Asn Lys Tyr Phe Val Thr Ile Gly
225                 230                 235                 240
Gly Phe Thr Gly Ala Asp Ala Ala Lys Asn Gly Asp Tyr Glu Val Asn
                245                 250                 255
Val Ala Thr Asp Gly Thr Val Thr Leu Ala Ala Gly Ala Thr Lys Thr
            260                 265                 270
Thr Met Pro Ala Gly Ala Thr Thr Lys Thr Glu Val Gln Glu Leu Lys
        275                 280                 285
Asp Thr Pro Ala Val Val Ser Ala Asp Ala Lys Asn Ala Leu Ile Ala
    290                 295                 300
Gly Gly Val Asp Ala Thr Asp Ala Asn Gly Ala Glu Leu Val Lys Met
305                 310                 315                 320
Ser Tyr Thr Asp Lys Asn Gly Lys Thr Ile Glu Gly Gly Tyr Ala Leu
                325                 330                 335
Lys Ala Gly Asp Lys Tyr Tyr Ala Ala Asp Tyr Asp Glu Ala Thr Gly
            340                 345                 350
Ala Ile Lys Ala Lys Thr Thr Ser Tyr Thr Ala Ala Asp Gly Thr Thr
        355                 360                 365
Lys Thr Ala Ala Asn Gln Leu Gly Gly Val Asp Gly Lys Thr Glu Val
    370                 375                 380
Val Thr Ile Asp Gly Lys Thr Tyr Asn Ala Ser Lys Ala Ala Gly His
385                 390                 395                 400
Asp Phe Lys Ala Gln Pro Glu Leu Ala Glu Ala Ala Ala Lys Thr Thr
                405                 410                 415
Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala
            420                 425                 430
Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile
        435                 440                 445
Thr Asn Leu Gly Asn Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg
    450                 455                 460
Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala
465                 470                 475                 480
Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln
                485                 490                 495
Val Pro Gln Asn Val Leu Ser Leu Leu Ala Lys Gly Thr Arg Thr Arg
            500                 505                 510
Gly Lys Leu Cys Pro Asp Cys Leu Asn Cys Thr Asp Leu Asp Val Ala
        515                 520                 525
Leu Gly Arg Pro Met Cys Val Gly Thr Thr Pro Ser Ala Lys Ala Ser
    530                 535                 540
Ile Leu His Glu Val Arg Pro Val Thr Ser Gly Cys Phe Pro Ile Met
545                 550                 555                 560
His Asp Arg Thr Lys Ile Arg Gln Leu Pro Asn Leu Leu Arg Gly Tyr
                565                 570                 575
Glu Asn Ile Arg Leu Ser Thr Gln Asn Val Ile Asp Ala Glu Lys Ala
```

580 585 590
Leu Gly Gly Pro Tyr Arg Leu Gly Thr Ser Gly Ser Cys Pro Asn Ala
595 600 605
Thr Ser Lys Ser Gly Phe Phe Ala Thr Met Ala Trp Ala Val Pro Lys
610 615 620
Asp Asn Asn Lys Asn Ala Thr Asn Pro Leu Thr Val Glu Val Pro Tyr
625 630 635 640
Ile Cys Thr Glu Gly Glu Asp Gln Ile Thr Val Trp Gly Phe His Ser
645 650 655
Asp Asp Lys Thr Gln Met Lys Asn Leu Tyr Gly Asp Ser Asn Pro Gln
660 665 670
Lys Phe Thr Ser Ser Ala Asn Gly Val Thr Thr His Tyr Val Ser Gln
675 680 685
Ile Gly Gly Phe Pro Asp Gln Thr Glu Asp Gly Gly Leu Pro Gln Ser
690 695 700
Gly Arg Ile Val Val Asp Tyr Met Val Gln Lys Pro Gly Lys Thr Gly
705 710 715 720
Thr Ile Val Tyr Gln Arg Gly Val Leu Leu Pro Gln Lys Val Trp Cys
725 730 735
Ala Ser Gly Arg Ser Lys Val Ile Lys Gly
740 745

<210> 212
<211> 1518
<212> PRT
<213> Artificial Sequence

<220>
<223> STF2

<400> 212
Ala Thr Gly Gly Cys Ala Cys Ala Ala Gly Thr Ala Ala Thr Cys Ala
1 5 10 15
Ala Cys Ala Cys Thr Ala Ala Cys Ala Gly Thr Cys Thr Gly Thr Cys
20 25 30
Gly Cys Thr Gly Cys Thr Gly Ala Cys Cys Cys Ala Gly Ala Ala Thr
35 40 45
Ala Ala Cys Cys Thr Gly Ala Ala Cys Ala Ala Ala Thr Cys Cys Cys
50 55 60
Ala Gly Thr Cys Cys Gly Cys Ala Cys Thr Gly Gly Gly Cys Ala Cys
65 70 75 80
Cys Gly Cys Thr Ala Thr Cys Gly Ala Gly Cys Gly Thr Cys Thr Gly
85 90 95
Thr Cys Thr Thr Cys Thr Gly Gly Thr Cys Thr Gly Cys Gly Thr Ala
100 105 110
Thr Cys Ala Ala Cys Ala Gly Cys Gly Cys Gly Ala Ala Ala Gly Ala
115 120 125
Cys Gly Ala Thr Gly Cys Gly Gly Cys Ala Gly Gly Thr Cys Ala Gly
130 135 140
Gly Cys Gly Ala Thr Thr Gly Cys Thr Ala Ala Cys Cys Gly Thr Thr
145 150 155 160
Thr Cys Ala Cys Cys Gly Cys Gly Ala Ala Cys Ala Thr Cys Ala Ala
165 170 175
Ala Gly Gly Thr Cys Thr Gly Ala Cys Thr Cys Ala Gly Gly Cys Thr
180 185 190
Thr Cys Cys Cys Gly Thr Ala Ala Cys Gly Cys Thr Ala Ala Cys Gly
195 200 205
Ala Cys Gly Gly Thr Ala Thr Cys Thr Cys Cys Ala Thr Thr Gly Cys

```
    210                 215                 220
Gly Cys Ala Gly Ala Cys Cys Ala Cys Thr Gly Ala Ala Gly Gly Cys
225                 230                 235                 240
Gly Cys Gly Cys Thr Gly Ala Ala Cys Gly Ala Ala Ala Thr Cys Ala
                245                 250                 255
Ala Cys Ala Ala Cys Ala Ala Cys Cys Thr Gly Cys Ala Gly Cys Gly
            260                 265                 270
Thr Gly Thr Gly Cys Gly Thr Gly Ala Ala Cys Thr Gly Gly Cys Gly
        275                 280                 285
Gly Thr Thr Cys Ala Gly Thr Cys Thr Gly Cys Thr Ala Ala Cys Ala
    290                 295                 300
Gly Cys Ala Cys Cys Ala Ala Cys Thr Cys Cys Cys Ala Gly Thr Cys
305                 310                 315                 320
Thr Gly Ala Cys Cys Thr Cys Gly Ala Cys Thr Cys Cys Ala Thr Cys
                325                 330                 335
Cys Ala Gly Gly Cys Thr Gly Ala Ala Ala Thr Cys Ala Cys Cys Cys
            340                 345                 350
Ala Gly Cys Gly Cys Cys Thr Gly Ala Ala Cys Gly Ala Ala Ala Thr
        355                 360                 365
Cys Gly Ala Cys Cys Gly Thr Gly Thr Ala Thr Cys Cys Gly Gly Cys
    370                 375                 380
Cys Ala Gly Ala Cys Thr Cys Ala Gly Thr Thr Cys Ala Ala Cys Gly
385                 390                 395                 400
Gly Cys Gly Thr Gly Ala Ala Ala Gly Thr Cys Cys Thr Gly Gly Cys
                405                 410                 415
Gly Cys Ala Gly Gly Ala Cys Ala Ala Cys Ala Cys Cys Cys Thr Gly
            420                 425                 430
Ala Cys Cys Ala Thr Cys Cys Ala Gly Gly Thr Thr Gly Gly Cys Gly
        435                 440                 445
Cys Cys Ala Ala Cys Gly Ala Cys Gly Gly Thr Gly Ala Ala Ala Cys
    450                 455                 460
Thr Ala Thr Cys Gly Ala Thr Ala Thr Cys Gly Ala Thr Cys Thr Gly
465                 470                 475                 480
Ala Ala Gly Cys Ala Gly Ala Thr Cys Ala Ala Cys Thr Cys Thr Cys
                485                 490                 495
Ala Gly Ala Cys Cys Cys Thr Gly Gly Gly Thr Cys Thr Gly Gly Ala
            500                 505                 510
Cys Thr Cys Ala Cys Thr Gly Ala Ala Cys Gly Thr Gly Cys Ala Gly
        515                 520                 525
Ala Ala Ala Gly Cys Gly Thr Ala Thr Gly Ala Thr Gly Thr Gly Ala
    530                 535                 540
Ala Ala Gly Ala Thr Ala Cys Ala Gly Cys Ala Gly Thr Ala Ala Cys
545                 550                 555                 560
Ala Ala Cys Gly Ala Ala Ala Gly Cys Thr Thr Ala Thr Gly Cys Cys
                565                 570                 575
Ala Ala Thr Ala Ala Thr Gly Gly Thr Ala Cys Thr Ala Cys Ala Cys
            580                 585                 590
Thr Gly Gly Ala Thr Gly Thr Ala Thr Cys Gly Gly Gly Thr Cys Thr
        595                 600                 605
Thr Gly Ala Thr Gly Ala Thr Gly Cys Ala Gly Cys Thr Ala Thr Thr
    610                 615                 620
Ala Ala Ala Gly Cys Gly Gly Cys Thr Ala Cys Gly Gly Gly Thr Gly
625                 630                 635                 640
Gly Thr Ala Cys Gly Ala Ala Thr Gly Gly Thr Ala Cys Gly Gly Cys
                645                 650                 655
Thr Thr Cys Thr Gly Thr Ala Ala Cys Cys Gly Gly Thr Gly Gly Thr
            660                 665                 670
Gly Cys Gly Gly Thr Thr Ala Ala Ala Thr Thr Thr Gly Ala Cys Gly
        675                 680                 685
```

```
Cys Ala Gly Ala Thr Ala Ala Thr Ala Ala Cys Ala Ala Gly Thr Ala
    690                 695                 700
Cys Thr Thr Thr Gly Thr Thr Ala Cys Thr Ala Thr Thr Gly Gly Thr
705                 710                 715                 720
Gly Gly Cys Thr Thr Thr Ala Cys Thr Gly Gly Thr Gly Cys Thr Gly
                725                 730                 735
Ala Thr Gly Cys Cys Gly Cys Cys Ala Ala Ala Ala Ala Thr Gly Gly
            740                 745                 750
Cys Gly Ala Thr Thr Ala Thr Gly Ala Ala Gly Thr Thr Ala Ala Cys
        755                 760                 765
Gly Thr Thr Gly Cys Thr Ala Cys Thr Gly Ala Cys Gly Gly Thr Ala
    770                 775                 780
Cys Ala Gly Thr Ala Ala Cys Cys Cys Thr Thr Gly Cys Gly Gly Cys
785                 790                 795                 800
Thr Gly Gly Cys Gly Cys Ala Ala Cys Thr Ala Ala Ala Ala Cys Cys
                805                 810                 815
Ala Cys Ala Ala Thr Gly Cys Cys Thr Gly Cys Thr Gly Gly Thr Gly
            820                 825                 830
Cys Gly Ala Cys Ala Ala Cys Thr Ala Ala Ala Ala Cys Ala Gly Ala
        835                 840                 845
Ala Gly Thr Ala Cys Ala Gly Gly Ala Gly Thr Thr Ala Ala Ala Ala
    850                 855                 860
Gly Ala Thr Ala Cys Ala Cys Cys Gly Gly Cys Ala Gly Thr Thr Gly
865                 870                 875                 880
Thr Thr Thr Cys Ala Gly Cys Ala Gly Ala Thr Gly Cys Thr Ala Ala
                885                 890                 895
Ala Ala Ala Thr Gly Cys Cys Thr Thr Ala Ala Thr Thr Gly Cys Thr
            900                 905                 910
Gly Gly Cys Gly Gly Cys Gly Thr Thr Gly Ala Cys Gly Cys Thr Ala
        915                 920                 925
Cys Cys Gly Ala Thr Gly Cys Thr Ala Ala Thr Gly Gly Cys Gly Cys
    930                 935                 940
Thr Gly Ala Gly Thr Thr Gly Gly Thr Cys Ala Ala Ala Ala Thr Gly
945                 950                 955                 960
Thr Cys Thr Thr Ala Thr Ala Cys Cys Gly Ala Thr Ala Ala Ala Ala
                965                 970                 975
Ala Thr Gly Gly Thr Ala Ala Gly Ala Cys Ala Ala Thr Thr Gly Ala
            980                 985                 990
Ala Gly Gly Cys Gly Gly Thr Thr Ala Thr Gly Cys Gly Cys Thr Thr
        995                 1000                1005
Ala Ala Ala Gly Cys Thr Gly Gly Cys Gly Ala Thr Ala Ala Gly Thr
    1010                1015                1020
Ala Thr Thr Ala Cys Gly Cys Cys Gly Cys Ala Gly Ala Thr Thr Ala
1025                1030                1035                1040
Cys Gly Ala Thr Gly Ala Ala Gly Cys Gly Ala Cys Ala Gly Gly Ala
                1045                1050                1055
Gly Cys Ala Ala Thr Thr Ala Ala Ala Gly Cys Thr Ala Ala Ala Ala
            1060                1065                1070
Cys Thr Ala Cys Ala Ala Ala Gly Thr Thr Ala Thr Ala Cys Thr Gly Cys
        1075                1080                1085
Thr Gly Cys Thr Gly Ala Cys Gly Gly Cys Ala Cys Thr Ala Cys Cys
    1090                1095                1100
Ala Ala Ala Ala Cys Ala Gly Cys Gly Gly Cys Thr Ala Ala Cys Cys
1105                1110                1115                1120
Ala Ala Cys Thr Gly Gly Gly Thr Gly Gly Cys Gly Thr Ala Gly Ala
                1125                1130                1135
Cys Gly Gly Thr Ala Ala Ala Ala Cys Cys Gly Ala Ala Gly Thr Cys
            1140                1145                1150
Gly Thr Thr Ala Cys Thr Ala Thr Cys Gly Ala Cys Gly Gly Thr Ala
```

```
        1155                1160                1165
Ala Ala Ala Cys Cys Thr Ala Cys Ala Ala Thr Gly Cys Cys Ala Gly
    1170                1175                1180
Cys Ala Ala Ala Gly Cys Cys Gly Cys Thr Gly Gly Thr Cys Ala Thr
1185                1190                1195                1200
Gly Ala Thr Thr Thr Cys Ala Ala Ala Gly Cys Ala Cys Ala Ala Cys
                1205                1210                1215
Cys Ala Gly Ala Gly Cys Thr Gly Gly Cys Gly Gly Ala Ala Gly Cys
            1220                1225                1230
Ala Gly Cys Cys Gly Cys Thr Ala Ala Ala Ala Cys Cys Ala Cys Cys
        1235                1240                1245
Gly Ala Ala Ala Ala Cys Cys Cys Gly Cys Thr Gly Cys Ala Gly Ala
    1250                1255                1260
Ala Ala Ala Thr Thr Gly Ala Thr Gly Cys Cys Gly Cys Gly Cys Thr
1265                1270                1275                1280
Gly Gly Cys Gly Cys Ala Gly Gly Thr Gly Gly Ala Thr Gly Cys Gly
                1285                1290                1295
Cys Thr Gly Cys Gly Cys Thr Cys Thr Gly Ala Thr Cys Thr Gly Gly
            1300                1305                1310
Gly Thr Gly Cys Gly Gly Thr Ala Cys Ala Ala Ala Ala Cys Cys Gly
        1315                1320                1325
Thr Thr Thr Cys Ala Ala Cys Thr Cys Thr Gly Cys Thr Ala Thr Cys
    1330                1335                1340
Ala Cys Cys Ala Ala Cys Cys Thr Gly Gly Gly Cys Ala Ala Thr Ala
1345                1350                1355                1360
Cys Cys Gly Thr Ala Ala Ala Cys Ala Ala Thr Cys Thr Gly Thr Cys
                1365                1370                1375
Thr Gly Ala Ala Gly Cys Gly Cys Gly Thr Ala Gly Cys Cys Gly Thr
            1380                1385                1390
Ala Thr Cys Gly Ala Ala Gly Ala Thr Thr Cys Cys Gly Ala Cys Thr
        1395                1400                1405
Ala Cys Gly Cys Gly Ala Cys Cys Gly Ala Ala Gly Thr Thr Thr Cys
    1410                1415                1420
Cys Ala Ala Cys Ala Thr Gly Thr Cys Thr Cys Gly Cys Gly Cys Gly
1425                1430                1435                1440
Cys Ala Gly Ala Thr Thr Cys Thr Gly Cys Ala Gly Cys Ala Gly Gly
                1445                1450                1455
Cys Cys Gly Gly Thr Ala Cys Thr Thr Cys Cys Gly Thr Thr Cys Thr
            1460                1465                1470
Gly Gly Cys Gly Cys Ala Gly Gly Cys Thr Ala Ala Cys Cys Ala Gly
        1475                1480                1485
Gly Thr Cys Cys Cys Gly Cys Ala Gly Ala Ala Cys Gly Thr Gly Cys
    1490                1495                1500
Thr Gly Thr Cys Thr Cys Thr Gly Thr Thr Ala Cys Gly Thr
1505                1510                1515


<210> 213
<211> 583
<212> PRT
<213> Influenza B/Yamagata/16/18 HA

<400> 213
Met Lys Ala Ile Ile Val Leu Leu Met Val Val Thr Ser Asn Ala Asp
 1               5                  10                  15
Arg Ile Cys Thr Gly Ile Thr Ser Ser Asn Ser Pro His Val Val Lys
            20                  25                  30
Thr Ala Thr Gln Gly Glu Val Asn Val Thr Gly Val Ile Pro Leu Thr
        35                  40                  45
```

```
Thr Thr Pro Thr Lys Ser His Phe Ala Asn Leu Lys Gly Thr Lys Thr
    50                  55                  60
Arg Gly Lys Leu Cys Pro Asn Cys Leu Asn Cys Thr Asp Leu Asp Val
65                  70                  75                  80
Ala Leu Gly Arg Pro Met Cys Met Gly Thr Ile Pro Ser Ala Lys Ala
                85                  90                  95
Ser Ile Leu His Glu Val Arg Pro Val Thr Ser Gly Cys Phe Pro Ile
            100                 105                 110
Met His Asp Arg Thr Lys Ile Arg Gln Leu Pro Asn Leu Leu Arg Gly
        115                 120                 125
Tyr Glu Asn Ile Arg Leu Ser Thr His Asn Val Ile Asn Ala Glu Arg
    130                 135                 140
Ala Pro Gly Gly Pro Tyr Arg Leu Gly Thr Ser Gly Ser Cys Pro Asn
145                 150                 155                 160
Val Thr Ser Arg Asn Gly Phe Phe Ala Thr Met Ala Trp Ala Val Pro
                165                 170                 175
Arg Asp Asn Lys Thr Ala Thr Asn Pro Leu Thr Val Glu Val Pro Tyr
            180                 185                 190
Ile Cys Thr Lys Gly Glu Asp Gln Ile Thr Val Trp Gly Phe His Ser
        195                 200                 205
Asp Asp Lys Thr Gln Met Lys Asn Leu Tyr Gly Asp Ser Asn Pro Gln
    210                 215                 220
Lys Phe Thr Ser Ser Ala Asn Gly Val Thr Thr His Tyr Val Ser Gln
225                 230                 235                 240
Ile Gly Asp Phe Pro Asn Gln Thr Glu Asp Gly Gly Leu Pro Gln Ser
                245                 250                 255
Gly Arg Ile Val Val Asp Tyr Met Val Gln Lys Pro Gly Lys Thr Gly
            260                 265                 270
Thr Ile Val Tyr Gln Arg Gly Val Leu Leu Pro Gln Lys Val Trp Cys
        275                 280                 285
Ala Ser Gly Arg Ser Lys Val Ile Lys Gly Ser Leu Pro Leu Ile Gly
    290                 295                 300
Glu Ala Asp Cys Leu His Glu Lys Tyr Gly Gly Leu Asn Lys Ser Lys
305                 310                 315                 320
Pro Tyr Tyr Thr Gly Glu His Ala Lys Ala Ile Gly Asn Cys Pro Ile
                325                 330                 335
Trp Val Lys Thr Pro Leu Lys Leu Ala Asn Gly Thr Lys Tyr Arg Pro
            340                 345                 350
Pro Ala Lys Leu Leu Lys Glu Arg Gly Phe Phe Gly Ala Ile Ala Gly
        355                 360                 365
Phe Leu Glu Gly Gly Trp Glu Gly Met Ile Ala Gly Trp His Gly Tyr
    370                 375                 380
Thr Ser His Gly Ala His Gly Val Ala Val Ala Ala Asp Leu Lys Ser
385                 390                 395                 400
Thr Gln Glu Ala Ile Asn Lys Ile Thr Lys Asn Leu Asn Ser Leu Ser
                405                 410                 415
Glu Leu Glu Val Lys Asn Leu Gln Arg Leu Ser Gly Ala Met Asp Glu
            420                 425                 430
Leu His Asn Glu Ile Leu Glu Leu Asp Glu Lys Val Asp Asp Leu Arg
        435                 440                 445
Ala Asp Thr Ile Ser Ser Gln Ile Glu Leu Ala Val Leu Leu Ser Asn
    450                 455                 460
Glu Gly Ile Ile Asn Ser Glu Asp Glu His Leu Leu Ala Leu Glu Arg
465                 470                 475                 480
Lys Leu Lys Lys Met Leu Gly Pro Ser Ala Val Asp Ile Gly Asn Gly
                485                 490                 495
Cys Phe Glu Thr Lys His Lys Cys Asn Gln Thr Cys Leu Asp Arg Ile
            500                 505                 510
Ala Ala Gly Thr Phe Asn Ala Gly Glu Phe Ser Leu Pro Thr Phe Asp
```

```
        515                 520                 525
Ser Leu Asn Ile Thr Ala Ala Ser Leu Asn Asp Asp Gly Leu Asp Asn
    530                 535                 540
His Thr Ile Leu Leu Tyr Tyr Ser Thr Ala Ala Ser Ser Leu Ala Val
545                 550                 555                 560
Thr Leu Met Ile Ala Ile Phe Ile Val Tyr Met Val Ser Arg Asp Asn
                565                 570                 575
Val Ser Cys Ser Ile Cys Leu
            580
```

<210> 214
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> PR/8 HA Peptide

<400> 214
Thr His Ser Val Asn Leu
 1   5

<210> 215
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> PR/8 HA Peptide

<400> 215
Ser His Asn Gly Lys Leu
 1   5

<210> 216
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> PR/8 HA Peptide

<400> 216
Cys Arg Leu Lys
 1

<210> 217
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> PR/8 HA Peptide

<400> 217

Pro Tyr Gln Asn
1

<210> 218
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> PR/8 HA Peptide

<400> 218
Pro Lys Tyr Val
1

<210> 219
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> PR/8 HA Peptide

<400> 219
Arg Ser Ala Lys
1

<210> 220
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> PR/8 HA Peptide

<400> 220
Asn Ala Ser Met His Glu
1 5

<210> 221
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> PR/8 HA Peptide

<400> 221
Ala Pro Leu Gln Leu Gly
1 5

<210> 222
<211> 5
<212> PRT

<213> Artificial Sequence

<220>
<223> PR/8 HA Peptide

<400> 222
Gly Ile Ile Thr Ser
1 5

<210> 223
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> PR/8 HA Peptide

<400> 223
Cys Asn Thr Lys
1

<210> 224
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> PR/8 HA Peptide

<400> 224
Ser Tyr Ile Val Glu Thr
1 5

<210> 225
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> PR/8 HA Peptide

<400> 225
Asn Ile Ala Gly Trp Leu Leu Gly
1 5

<210> 226
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> PR/8 HA Peptide

<400> 226
Pro Gly Asp Phe Ile

1 5

<210> 227
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> PR/8 HA Peptide

<400> 227
Asn Ser Glu Asn Gly Ile Cys Tyr
1 5

<210> 228
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> PR/8 HA Peptide

<400> 228
Ala Phe Ala Leu Ser Arg Gly Phe
1 5

<210> 229
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> PR/8 HA Peptide

<400> 229
Ala Phe Ala Leu Ser Arg Gly
1 5

<210> 230
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> VN04 HA peptide

<400> 230
Thr His Ala Gln Asp Ile
1 5

<210> 231
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> VN04 HA peptide

<400> 231
Glu Lys Lys His Asn Gly
1 5

<210> 232
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> VN04 HA peptide

<400> 232
Lys Leu Cys Asp Leu Asp
1 5

<210> 233
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> VN04 HA peptide

<400> 233
Lys Ser Asn Arg
1

<210> 234
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> VN04 HA peptide

<400> 234
Pro Lys Tyr Val
1

<210> 235
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> VN04 HA peptide

<400> 235
Leu Val Leu Ala Thr Gly
1 5

<210> 236
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> VN04 HA peptide

<400> 236
Lys Pro Leu Ile Leu
1 5

<210> 237
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> VN04 HA peptide

<400> 237
Thr Ile Met Lys Ser
1 5

<210> 238
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> VN04 HA peptide

<400> 238
Tyr Gly Asn Cys Asn
1 5

<210> 239
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> VN04 HA peptide

<400> 239
Ser Tyr Ile Val Glu Lys
1 5

<210> 240
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> VN04 HA peptide

<400> 240
Pro Gly Asp Phe Asn
1 5

<210> 241
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> VN04 HA peptide

<400> 241
Lys Gly Asp Ser
1

<210> 242
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> VN04 HA peptide

<400> 242
Glu Tyr Ala Tyr Lys Ile Val Lys
1 5

<210> 243
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> IND05 HA peptide

<400> 243
Thr His Ala Gln Asp Ile
1 5

<210> 244
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> IND05 HA peptide

<400> 244
Glu Lys Thr His Asn Gly
1 5

<210> 245
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> IND05 HA peptide

<400> 245
Lys Leu Cys Asp Leu Asp
1 5

<210> 246
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> IND05 HA peptide

<400> 246
Lys Ser Asn Arg
1

<210> 247
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> IND05 HA peptide

<400> 247
Pro Lys Tyr Val
1

<210> 248
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> IND05 HA peptide

<400> 248
Leu Val Leu Ala Thr Gly
1 5

<210> 249
<211> 5
<212> PRT
<213> Artificial Sequence

<220>

<223> IND05 HA peptide

<400> 249
Lys Pro Leu Ile Leu
1 5

<210> 250
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> IND05 HA peptide

<400> 250
Ala Ile Met Lys Ser
1 5

<210> 251
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> IND05 HA peptide

<400> 251
Tyr Gly Asn Cys Asn
1 5

<210> 252
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> IND05 HA peptide

<400> 252
Ser Tyr Ile Val Glu Lys
1 5

<210> 253
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> IND05 HA peptide

<400> 253
Pro Gly Ser Phe Asn
1 5

<210> 254
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> IND05 HA peptide

<400> 254
Glu Tyr Ala Tyr Lys Ile Val Lys
1 5

<210> 255
<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> New Caledonia HA peptide

<400> 255
Thr His Ser Val Asn Leu Leu Glu Asp
1 5

<210> 256
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> New Caledonia HA peptide

<400> 256
Ser His Asn Gly Lys Leu
1 5

<210> 257
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> New Caledonia HA peptide

<400> 257
Ala Pro Leu Gln Leu Gly
1 5

<210> 258
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> New Caledonia HA peptide

<400> 258
Arg Ser Ala Lys
1

<210> 259
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> New Caledonia HA peptide

<400> 259
Pro Lys Tyr Val
1

<210> 260
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> New Caledonia HA peptide

<400> 260
Leu Arg Met Val Thr Gly
1 5

<210> 261
<211> 12
<212> PRT
<213> Artificial Sequence

<220>
<223> New Caledonia HA peptide

<400> 261
Ser His Asn Gly Lys Leu Cys Leu Leu Lys Gly Ile
1 5 10

<210> 262
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> New Caledonia HA peptide

<400> 262
Ser Gly Ile Ile Thr Ser
1 5

<210> 263

<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> New Caledonia HA peptide

<400> 263
Asn Ala Pro Met Asp Glu Cys Asp Ala
1 5

<210> 264
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> New Caledonia HA peptide

<400> 264
Ser Tyr Ile Val Glu Thr
1 5

<210> 265
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> New Caledonia HA peptide

<400> 265
Pro Gly Tyr Phe Ala
1 5

<210> 266
<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> New Caledonia HA peptide

<400> 266
Tyr Ala Phe Ala Leu Ser Arg Gly Phe
1 5

<210> 267
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> A/Wisconsin/67/2005 HA Peptide

<400> 267
Thr Asn Ala Thr Glu
1 5

<210> 268
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> A/Wisconsin/67/2005 HA Peptide

<400> 268
Pro Arg Tyr Val Lys
1 5

<210> 269
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> A/Wisconsin/67/2005 HA Peptide

<400> 269
Pro Phe Gln Asn Val
1 5

<210> 270
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> A/Wisconsin/67/2005 HA Peptide

<400> 270
Gln Ile Leu Asp Gly Glu
1 5

<210> 271
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> A/Wisconsin/67/2005 HA Peptide

<400> 271
Gly Gly Ile Cys Asp
1 5

<210> 272
<211> 6

<212> PRT
<213> Artificial Sequence

<220>
<223> A/Wisconsin/67/2005 HA Peptide

<400> 272
Ser Ser Ile Met Arg Ser
1 5

<210> 273
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> A/Wisconsin/67/2005 HA Peptide

<400> 273
Ala Pro Ile Gly Lys
1 5

<210> 274
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> A/Wisconsin/67/2005 HA Peptide

<400> 274
Leu Phe Val Glu Arg
1 5

<210> 275
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> A/Wisconsin/67/2005 HA Peptide

<400> 275
Thr Leu Ile Asp Ala Leu Leu Gly
1 5

<210> 276
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> A/Wisconsin/67/2005 HA Peptide

<400> 276

Arg Gly Tyr Phe Lys Ile
1 5

<210> 277
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> PR/8 HA1-3mut peptide

<400> 277
Asn Ser Glu Asn Glu Ile Cys Tyr
1 5

<210> 278
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> PR/8 HA1-3mut peptide

<400> 278
Lys Glu Glu Leu Arg Glu Gln Leu
1 5

<210> 279
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> B/Lee/40 HA1-1 peptide

<400> 279
Thr Thr Thr Pro Thr Lys
1 5

<210> 280
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> B/Lee/40 HA1-1 peptide

<400> 280
Cys Pro Ile Trp Val Lys
1 5

<210> 281
<211> 6
<212> PRT

<213> Artificial Sequence

<220>
<223> B/Lee/40 HA1-2 peptide

<400> 281
Lys Gly Thr Gln Thr Arg
1 5

<210> 282
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> B/Lee/40 HA1-2 peptide

<400> 282
Ser Lys Val Ile Lys Gly
1 5

<210> 283
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> A/Aichi/2/68 HA1-1 peptide

<400> 283
Gln Ser Ser Ser Thr Gly
1 5

<210> 284
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> A/Aichi/2/68 HA1-1 peptide

<400> 284
Cys Pro Lys Tyr Val Lys
1 5

<210> 285
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> A/Aichi/2/68 HA1-2 peptide

<400> 285
His Arg Ile Leu Asp Gly

1 5

<210> 286
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> A/Aichi/2/68 HA1-2 peptide

<400> 286
Ser Ile Met Arg Ser Asp
1 5

<210> 287
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> B/Malaysia/2506/2004 HA1-1 peptide

<400> 287
Thr Thr Thr Pro Thr Lys
1 5

<210> 288
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> B/Malaysia/2506/2004 HA1-1 peptide

<400> 288
Cys Pro Ile Trp Val Lys
1 5

<210> 289
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> B/Malaysia/2506/2004 HA1-2 peptide

<400> 289
Lys Gly Thr Glu Thr Arg
1 5

<210> 290
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> B/Malaysia/2506/2004 HA1-2 peptide

<400> 290
Ser Lys Val Ile Lys Gly
1 5

<210> 291
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> B/Ohio/1/2005 HA1-1 peptide

<400> 291
Thr Thr Thr Pro Thr Lys
1 5

<210> 292
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> B/Ohio/1/2005 HA1-1 peptide

<400> 292
Cys Pro Ile Trp Val Lys
1 5

<210> 293
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> B/Ohio/1/2005 HA1-2 peptide

<400> 293
Lys Gly Thr Lys Thr Arg
1 5

<210> 294
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> B/Ohio/1/2005 HA1-2 peptide

<400> 294
Ser Lys Val Ile Lys Gly
1 5

<210> 295
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> B/Shanghai/361/2002 HA1-1 peptide

<400> 295
Thr Thr Thr Pro Ile Lys
1 5

<210> 296
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> B/Shanghai/361/2002 HA1-1 peptide

<400> 296
Cys Pro Ile Trp Val Lys
1 5

<210> 297
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> B/Shanghai/361/2002 HA1-2 peptide

<400> 297
Lys Gly Thr Arg Thr Arg
1 5

<210> 298
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> B/Shanghai/361/2002 HA1-2 peptide

<400> 298
Ser Lys Val Ile Lys Gly
1 5

<210> 299
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> B/Lee/40 HA peptide

<400> 299
Thr Gly Val Ile
1

<210> 300
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> B/Lee/40 HA peptide

<400> 300
Thr Thr Thr Pro
1

<210> 301
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> B/Lee/40 HA peptide

<400> 301
Pro Ile Trp Val
1

<210> 302
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> B/Lee/40 HA peptide

<400> 302
Pro Tyr Tyr Thr Gly
1 5

<210> 303
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> B/Lee/40 HA peptide

<400> 303
Asp Cys Leu His Glu
1 5

<210> 304
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> B/Lee/40 HA peptide

<400> 304
Tyr Gly Gly Leu Asn
1 5

<210> 305
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> B/Lee/40 HA peptide

<400> 305
Lys Ala Ile Gly Asn
1 5

<210> 306
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> B/Lee/40 HA peptide

<400> 306
Thr Gln Thr Arg Gly
1 5

<210> 307
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> B/Lee/40 HA peptide

<400> 307
Lys Val Ile Lys Gly
1 5

<210> 308
<211> 5
<212> PRT
<213> Artificial Sequence

<220>

<223> B/Lee/40 HA peptide

<400> 308
Asp Cys Leu His Glu
1 5

<210> 309
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> B/Lee/40 HA peptide

<400> 309
Ser Ile Leu His Glu Val
1 5

<210> 310
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2

<400> 310
gtgctgagcc tgttacgt 18

<210> 311
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> linker

<400> 311
Ser Gly Ser Gly Ser Gly Ser
1 5

<210> 312
<211> 506
<212> PRT
<213> S. typhimurium

<400> 312
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
1 5 10 15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
20 25 30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
35 40 45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
50 55 60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly

```
65                  70                  75                  80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
                85                  90                  95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
            100                 105                 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
        115                 120                 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
    130                 135                 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145                 150                 155                 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val Gln
                165                 170                 175
Lys Ala Tyr Asp Val Lys Asp Thr Ala Val Thr Thr Lys Ala Tyr Ala
            180                 185                 190
Asn Asn Gly Thr Thr Leu Asp Val Ser Gly Leu Asp Asp Ala Ala Ile
        195                 200                 205
Lys Ala Ala Thr Gly Gly Thr Asn Gly Thr Ala Ser Val Thr Gly Gly
    210                 215                 220
Ala Val Lys Phe Asp Ala Asp Asn Asn Lys Tyr Phe Val Thr Ile Gly
225                 230                 235                 240
Gly Phe Thr Gly Ala Asp Ala Ala Lys Asn Gly Asp Tyr Glu Val Asn
                245                 250                 255
Val Ala Thr Asp Gly Thr Val Thr Leu Ala Ala Gly Ala Thr Lys Thr
            260                 265                 270
Thr Met Pro Ala Gly Ala Thr Thr Lys Thr Glu Val Gln Glu Leu Lys
        275                 280                 285
Asp Thr Pro Ala Val Val Ser Ala Asp Ala Lys Asn Ala Leu Ile Ala
    290                 295                 300
Gly Gly Val Asp Ala Thr Asp Ala Asn Gly Ala Glu Leu Val Lys Met
305                 310                 315                 320
Ser Tyr Thr Asp Lys Asn Gly Lys Thr Ile Glu Gly Gly Tyr Ala Leu
                325                 330                 335
Lys Ala Gly Asp Lys Tyr Tyr Ala Ala Asp Tyr Asp Glu Ala Thr Gly
            340                 345                 350
Ala Ile Lys Ala Lys Thr Thr Ser Tyr Thr Ala Ala Asp Gly Thr Thr
        355                 360                 365
Lys Thr Ala Ala Asn Gln Leu Gly Gly Val Asp Gly Lys Thr Glu Val
    370                 375                 380
Val Thr Ile Asp Gly Lys Thr Tyr Asn Ala Ser Lys Ala Ala Gly His
385                 390                 395                 400
Asp Phe Lys Ala Gln Pro Glu Leu Ala Glu Ala Ala Ala Lys Thr Thr
                405                 410                 415
Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala
            420                 425                 430
Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile
        435                 440                 445
Thr Asn Leu Gly Asn Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg
    450                 455                 460
Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala
465                 470                 475                 480
Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln
                485                 490                 495
Val Pro Gln Asn Val Leu Ser Leu Leu Arg
            500                 505
```

<210> 313
<211> 277

<212> PRT
<213> Artificial Sequence

<220>
<223> STF2delta

<400> 313
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
1 5 10 15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
20 25 30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
35 40 45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
50 55 60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65 70 75 80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
85 90 95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
100 105 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
115 120 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
130 135 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145 150 155 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val His
165 170 175
Gly Ala Pro Val Asp Pro Ala Ser Pro Trp Thr Glu Asn Pro Leu Gln
180 185 190
Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala Leu Arg Ser Asp Leu
195 200 205
Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile Thr Asn Leu Gly Asn
210 215 220
Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg Ile Glu Asp Ser Asp
225 230 235 240
Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala Gln Ile Leu Gln Gln
245 250 255
Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln Val Pro Gln Asn Val
260 265 270
Leu Ser Leu Leu Arg
275

<210> 314
<211> 831
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2delta

<400> 314
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300

aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcatgg agcgccggtg 540
gatcctgcta gcccatggac cgaaaacccg ctgcagaaaa ttgatgccgc gctggcgcag 600
gtggatgcgc tgcgctctga tctgggtgcg gtacaaaacc gtttcaactc tgctatcacc 660
aacctgggca ataccgtaaa caatctgtct gaagcgcgta gccgtatcga agattccgac 720
tacgcgaccg aagtttccaa catgtctcgc gcgcagattt tgcagcaggc cggtacttcc 780
gttctggcgc aggctaacca ggtcccgcag aacgtgctgt ctctgttacg t 831

<210> 315
<211> 1518
<212> DNA
<213> S. typhimurium

<400> 315
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcagaa agcgtatgat 540
gtgaaagata cagcagtaac aacgaaagct tatgccaata atggtactac actggatgta 600
tcgggtcttg atgatgcagc tattaaagcg gctacgggtg gtacgaatgg tacggcttct 660
gtaaccggtg gtgcggttaa atttgacgca gataataaca agtactttgt tactattggt 720
ggctttactg gtgctgatgc cgccaaaaat ggcgattatg aagttaacgt tgctactgac 780
ggtacagtaa cccttgcggc tggcgcaact aaaaccacaa tgcctgctgg tgcgacaact 840
aaaacagaag tacaggagtt aaaagataca ccggcagttg tttcagcaga tgctaaaaat 900
gccttaattg ctggcggcgt tgacgctacc gatgctaatg gcgctgagtt ggtcaaaatg 960
tcttataccg ataaaaatgg taagacaatt gaaggcggtt atgcgcttaa agctggcgat 1020
aagtattacg ccgcagatta cgatgaagcg acaggagcaa ttaaagctaa aactacaagt 1080
tatactgctg ctgacggcac taccaaaaca gcggctaacc aactgggtgg cgtagacggt 1140
aaaaccgaag tcgttactat cgacggtaaa acctacaatg ccagcaaagc cgctggtcat 1200
gatttcaaag cacaaccaga gctggcggaa gcagccgcta aaaccaccga aaacccgctg 1260
cagaaaattg atgccgcgct ggcgcaggtg gatgcgctgc gctctgatct gggtgcggta 1320
caaaaccgtt tcaactctgc tatcaccaac ctgggcaata ccgtaaacaa tctgtctgaa 1380
gcgcgtagcc gtatcgaaga ttccgactac gcgaccgaag tttccaacat gtctcgcgcg 1440
cagattctgc agcaggccgg tacttccgtt ctggcgcagg ctaaccaggt cccgcagaac 1500
gtgctgtctc tgttacgt 1518

<210> 316
<211> 2763
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2.OVA

<400> 316
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420

gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcagaa agcgtatgat 540
gtgaaagata cagcagtaac aacgaaagct tatgccaata atggtactac actggatgta 600
tcgggtcttg atgatgcagc tattaaagcg gctacgggtg gtacgaatgg tacggcttct 660
gtaaccggtg gtgcggttaa atttgacgca gataataaca agtactttgt tactattggt 720
ggctttactg gtgctgatgc cgccaaaaat ggcgattatg aagttaacgt tgctactgac 780
ggtacagtaa cccttgcggc tggcgcaact aaaaccacaa tgcctgctgg tgcgacaact 840
aaaacagaag tacaggagtt aaaagataca ccggcagttg tttcagcaga tgctaaaaat 900
gccttaattg ctggcggcgt tgacgctacc gatgctaatg gcgctgagtt ggtcaaaatg 960
tcttataccg ataaaaatgg taagacaatt gaaggcggtt atgcgcttaa agctggcgat 1020
aagtattacg ccgcagatta cgatgaagcg acaggagcaa ttaaagctaa aaccacaagt 1080
tatactgctg ctgacggcac taccaaaaca gcggctaacc aactgggtgg cgtagacggt 1140
aaaaccgaag tcgttactat cgacggtaaa acctacaatg ccagcaaagc cgctggtcat 1200
gatttcaaag cacaaccaga gctggcggaa gcagccgcta aaaccaccga aaacccgctg 1260
cagaaaattg atgccgcgct ggcgcaggtg gatgcgctgc gctctgatct gggtgcggta 1320
caaaaccgtt tcaactctgc tatcaccaac ctgggcaata ccgtaaacaa tctgtctgaa 1380
gcgcgtagcc gtatcgaaga ttccgactac gcgaccgaag tttccaacat gtctcgcgcg 1440
cagattttgc agcaggccgg tacttccgtt ctggcgcagg ctaaccaggt cccgcagaac 1500
gtgctgtctc tgttacgtct cgagggctcc atcggcgcag caagcatgga attttgtttt 1560
gatgtattca aggagctcaa agtccaccat gccaatgaga acatcttcta ctgccccatt 1620
gccatcatgt cagctctagc catggtatac ctgggtgcaa aagacagcac caggacacaa 1680
ataaataagg ttgttcgctt tgataaactt ccaggattcg gagacagtat tgaagctcag 1740
tgtggcacat ctgtaaacgt tcactcttca cttagagaca tcctcaacca aatcaccaaa 1800
ccaaatgatg tttattcgtt cagccttgcc agtagacttt atgctgaaga gagataccca 1860
atcctgccag aatacttgca gtgtgtgaag gaactgtata gaggaggctt ggaacctatc 1920
aactttcaaa cagctgcaga tcaagccaga gagctcatca attcctgggt agaaagtcag 1980
acaaatggaa ttatcagaaa tgtccttcag ccaagctccg tggattctca aactgcaatg 2040
gttctggtta atgccattgt cttcaaagga ctgtgggaga aagcatttaa ggatgaagac 2100
acacaagcaa tgcctttcag agtgactgag caagaaagca aacctgtgca gatgatgtac 2160
cagattggtt tatttagagt ggcatcaatg gcttctgaga aaatgaagat cctggagctt 2220
ccatttgcca gtgggacaat gagcatgttg gtgctgttgc ctgatgaagt ctcaggcctt 2280
gagcagcttg agagtataat caactttgaa aaactgactg aatggaccag ttctaatgtt 2340
atggaagaga ggaagatcaa agtgtactta cctcgcatga agatggagga aaaatacaac 2400
ctcacatctg tcttaatggc tatgggcatt actgacgtgt ttagctcttc agccaatctg 2460
tctggcatct cctcagcaga gagcctgaag atatctcaag ctgtccatgc agcacatgca 2520
gaaatcaatg aagcaggcag agaggtggta gggtcagcag aggctggagt ggatgctgca 2580
agcgtctctg aagaatttag ggctgaccat ccattcctct tctgtatcaa gcacatcgca 2640
accaacgccg ttctcttctt tggcagatgt gtttcccctt cgaagcttga aggtaagcct 2700
atccctaacc ctctcctcgg tctcgattct acgcgtaccg gtcatcatca ccatcaccat 2760
tga 2763

<210> 317
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> primer STF28BGF-1

<400> 317
ctcgggagat ctgcacaagt aatcaacact aacagtct 38

<210> 318
<211> 41
<212> DNA
<213> Artificial Sequence

<220>
<223> primer STF28MCR-1

<400> 318
ccatgggcta gcaggatcca ccggcgctcc ctgcacgttc a 41

<210> 319
<211> 42
<212> DNA
<213> Artificial Sequence

<220>
<223> primer STF28MCF-2

<400> 319
ggagcgccgg tggatcctgc tagcccatgg accgaaaacc cg 42

<210> 320
<211> 84
<212> DNA
<213> Artificial Sequence

<220>
<223> primer STF28ECR-2

<400> 320
tctgcagaat tcacgtaaca gagacagcac gttctgcggg acgtcccgca gaacgtgctg 60
tctctgttac gtgaattctg caga 84

<210> 321
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> primer 3'Forward1

<400> 321
actgagtgca tatggcacaa gtaatcaaca ctaacag 37

<210> 322
<211> 55
<212> DNA
<213> Artificial Sequence

<220>
<223> primer 3'Reverse1

<400> 322
gactgactgc tcagcctatt agcagagcgg ccgccactgt gctggatatc tagag 55

<210> 323
<211> 44
<212> DNA
<213> Artificial Sequence

<220>
<223> primer 5'Forward2

<400> 323
agtcaggcca tatgtgcgca caagtaatca acactaacag tctg 44

<210> 324
<211> 59
<212> DNA
<213> Artificial Sequence

<220>
<223> primer 5"Reverse2

<400> 324
gactgactgc tcagcctatt aacgtaacag agacagcacg ttctgcggga cctggttag 59

<210> 325
<211> 870
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2delta.3'Cys

<400> 325
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcatgg agcgccggtg 540
gatcctgcta gcccatggac cgaaaacccg ctgcagaaaa ttgatgccgc gctggcgcag 600
gtggatgcgc tgcgctctga tctgggtgcg gtacaaaacc gtttcaactc tgctatcacc 660
aacctgggca ataccgtaaa caatctgtct gaagcgcgta gccgtatcga agattccgac 720
tacgcgaccg aagtttccaa catgtctcgc gcgcagattt tgcagcaggc cggtacttcc 780
gttctggcgc aggctaacca ggtcccgcag aacgtgctgt ctctgttacg tgaattctct 840
agatatccag cacagtggcg gccgctctgc 870

<210> 326
<211> 834
<212> DNA
<213> Artificial Sequence

<220>
<223> 5'Cys.STF2delta

<400> 326
atgtgcgcac aagtaatcaa cactaacagt ctgtcgctgc tgacccagaa taacctgaac 60
aaatcccagt ccgcactggg caccgctatc gagcgtctgt cttctggtct gcgtatcaac 120
agcgcgaaag acgatgcggc aggtcaggcg attgctaacc gtttcaccgc gaacatcaaa 180
ggtctgactc aggcttcccg taacgctaac gacggtatct ccattgcgca gaccactgaa 240
ggcgcgctga acgaaatcaa caacaacctg cagcgtgtgc gtgaactggc ggttcagtct 300
gctaacagca ccaactccca gtctgacctc gactccatcc aggctgaaat cacccagcgc 360
ctgaacgaaa tcgaccgtgt atccggccag actcagttca acggcgtgaa agtcctggcg 420
caggacaaca ccctgaccat ccaggttggc gccaacgacg gtgaaactat cgatatcgat 480
ctgaagcaga tcaactctca gaccctgggt ctggactcac tgaacgtgca tggagcgccg 540
gtggatcctg ctagcccatg gaccgaaaac ccgctgcaga aaattgatgc cgcgctggcg 600
caggtggatg cgctgcgctc tgatctgggt gcggtacaaa accgtttcaa ctctgctatc 660
accaacctgg gcaataccgt aaacaatctg tctgaagcgc gtagccgtat cgaagattcc 720
gactacgcga ccgaagtttc caacatgtct cgcgcgcaga ttttgcagca ggccggtact 780

tccgttctgg cgcaggctaa ccaggtcccg cagaacgtgc tgtctctgtt acgt 834

<210> 327
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer HingeForward1

<400> 327
actgagtgca tatggcacaa gtaatcaaca ctaacag 37

<210> 328
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer HingeReverse3

<400> 328
ggtccatggg caagcaggat ccaccggcgc t 31

<210> 329
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer HingeForward2

<400> 329
agcgccggtg gatcctgctt gcccatggac c 31

<210> 330
<211> 59
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer HingeReverse4

<400> 330
gactgactgc tcagcctatt aacgtaacag agacagcacg ttctgcggga cctggttag 59

<210> 331
<211> 831
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2delta.HingeCys

<400> 331
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240

gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcatgg agcgccggtg 540
gatcctgctt gcccatggac cgaaaacccg ctgcagaaaa ttgatgccgc gctggcgcag 600
gtggatgcgc tgcgctctga tctgggtgcg gtacaaaacc gtttcaactc tgctatcacc 660
aacctgggca ataccgtaaa caatctgtct gaagcgcgta gccgtatcga agattccgac 720
tacgcgaccg aagtttccaa catgtctcgc gcgcagattt tgcagcaggc cggtacttcc 780
gttctggcgc aggctaacca ggtcccgcag aacgtgctgt ctctgttacg t 831

<210> 332
<211> 290
<212> PRT
<213> Artificial Sequence

<220>
<223> STF2delta.3'Cys

<400> 332
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
1 5 10 15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
20 25 30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
35 40 45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
50 55 60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65 70 75 80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
85 90 95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
100 105 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
115 120 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
130 135 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145 150 155 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val His
165 170 175
Gly Ala Pro Val Asp Pro Ala Ser Pro Trp Thr Glu Asn Pro Leu Gln
180 185 190
Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala Leu Arg Ser Asp Leu
195 200 205
Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile Thr Asn Leu Gly Asn
210 215 220
Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg Ile Glu Asp Ser Asp
225 230 235 240
Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala Gln Ile Leu Gln Gln
245 250 255
Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln Val Pro Gln Asn Val
260 265 270
Leu Ser Leu Leu Arg Glu Phe Ser Arg Tyr Pro Ala Gln Trp Arg Pro
275 280 285
Leu Cys
290

<210> 333
<211> 277
<212> PRT
<213> Artificial Sequence

<220>
<223> STF2delta.HingeCys

<400> 333
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
1 5 10 15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
20 25 30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
35 40 45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
50 55 60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65 70 75 80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
85 90 95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
100 105 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
115 120 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
130 135 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145 150 155 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val His
165 170 175
Gly Ala Pro Val Asp Pro Ala Cys Pro Trp Thr Glu Asn Pro Leu Gln
180 185 190
Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala Leu Arg Ser Asp Leu
195 200 205
Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile Thr Asn Leu Gly Asn
210 215 220
Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg Ile Glu Asp Ser Asp
225 230 235 240
Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala Gln Ile Leu Gln Gln
245 250 255
Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln Val Pro Gln Asn Val
260 265 270
Leu Ser Leu Leu Arg
275

<210> 334
<211> 278
<212> PRT
<213> Artificial Sequence

<220>
<223> 5'Cys.STF2delta

<400> 334
Met Cys Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln
1 5 10 15

```
Asn Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg
            20                  25                  30
Leu Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly
        35                  40                  45
Gln Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln
    50                  55                  60
Ala Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu
65                  70                  75                  80
Gly Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu
                85                  90                  95
Ala Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser
            100                 105                 110
Ile Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser
        115                 120                 125
Gly Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr
    130                 135                 140
Leu Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp
145                 150                 155                 160
Leu Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val
                165                 170                 175
His Gly Ala Pro Val Asp Pro Ala Ser Pro Trp Thr Glu Asn Pro Leu
            180                 185                 190
Gln Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala Leu Arg Ser Asp
        195                 200                 205
Leu Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile Thr Asn Leu Gly
    210                 215                 220
Asn Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg Ile Glu Asp Ser
225                 230                 235                 240
Asp Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala Gln Ile Leu Gln
                245                 250                 255
Gln Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln Val Pro Gln Asn
            260                 265                 270
Val Leu Ser Leu Leu Arg
        275


<210> 335
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> lipopeptide

<400> 335
Cys Ser Lys Lys Lys Lys Gly Cys
 1               5


<210> 336
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> linker

<400> 336
Gly Ala Pro Val Asp Pro Ala Ser Pro Trp
```

1 5 10

<210> 337
<211> 19
<212> PRT
<213> Artificial Sequence

<220>
<223> H1C1 peptide

<400> 337
Asn Ile Pro Ser Ile Gln Ser Arg Gly Leu Phe Phe Ala Ile Ala Gly
1 5 10 15
Phe Ile Glu

<210> 338
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Cys H1C1

<400> 338
Cys Asn Ile Pro Ser Ile Gln Ser Arg Gly Leu Phe Phe Ala Ile Ala
1 5 10 15
Gly Phe Ile Glu
20

<210> 339
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Lys H1C1

<400> 339
Lys Asn Ile Pro Ser Ile Gln Ser Arg Gly Leu Phe Phe Ala Ile Ala
1 5 10 15
Gly Phe Ile Glu
20

<210> 340
<211> 1518
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2.blp

<400> 340
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120

gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcagaa agcgtatgat 540
gtgaaagata cagcagtaac aacgaaagct tatgccaata atggtactac actggatgta 600
tcgggtcttg atgatgcagc tattaaagcg gctacgggtg gtacgaatgg tacggcttct 660
gtaaccggtg gtgcggttaa atttgacgca gataataaca agtactttgt tactattggt 720
ggctttactg gtgctgatgc cgccaaaaat ggcgattatg aagttaacgt tgctactgac 780
ggtacagtaa cccttgcggc tggcgcaact aaaaccacaa tgcctgctgg tgcgacaact 840
aaaacagaag tacaggagtt aaaagataca ccggcagttg tttcagcaga tgctaaaaat 900
gccttaattg ctggcggcgt tgacgctacc gatgctaatg gcgctgagtt ggtcaaaatg 960
tcttataccg ataaaaatgg taagacaatt gaaggcggtt atgcgcttaa agctggcgat 1020
aagtattacg ccgcagatta cgatgaagcg acaggagcaa ttaaagctaa aactacaagt 1080
tatactgctg ctgacggcac taccaaaaca gcggctaacc aactgggtgg cgtagacggt 1140
aaaaccgaag tcgttactat cgacggtaaa acctacaatg ccagcaaagc cgctggtcat 1200
gatttcaaag cacaaccaga gctggcggaa gcagccgcta aaaccaccga aaacccgctg 1260
cagaaaattg atgccgcgct ggcgcaggtg gatgcgctgc gctctgatct gggtgcggta 1320
caaaaccgtt tcaactctgc tatcaccaac ctgggcaata ccgtaaacaa tctgtctgaa 1380
gcgcgtagcc gtatcgaaga ttccgactac gcgaccgaag tttccaacat gtctcgcgcg 1440
cagattctgc agcaggccgg tacttccgtt ctggcgcagg ctaaccaggt cccgcagaac 1500
gtgctgagcc tgttacgt 1518

<210> 341
<211> 19
<212> PRT
<213> Artificial Sequence

<220>
<223> HA maturation cleavage fragment (H1N1)

<400> 341
Asn Ile Pro Ser Ile Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly
1 5 10 15
Phe Ile Glu

<210> 342
<211> 529
<212> PRT
<213> Artificial Sequence

<220>
<223> STF2.1xH1Cl

<400> 342
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
1 5 10 15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
20 25 30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
35 40 45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
50 55 60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65 70 75 80

```
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
                85                  90                  95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
            100                 105                 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
        115                 120                 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
    130                 135                 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145                 150                 155                 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val Gln
                165                 170                 175
Lys Ala Tyr Asp Val Lys Asp Thr Ala Val Thr Thr Lys Ala Tyr Ala
            180                 185                 190
Asn Asn Gly Thr Thr Leu Asp Val Ser Gly Leu Asp Asp Ala Ala Ile
        195                 200                 205
Lys Ala Ala Thr Gly Gly Thr Asn Gly Thr Ala Ser Val Thr Gly Gly
    210                 215                 220
Ala Val Lys Phe Asp Ala Asp Asn Asn Lys Tyr Phe Val Thr Ile Gly
225                 230                 235                 240
Gly Phe Thr Gly Ala Asp Ala Ala Lys Asn Gly Asp Tyr Glu Val Asn
                245                 250                 255
Val Ala Thr Asp Gly Thr Val Thr Leu Ala Ala Gly Ala Thr Lys Thr
            260                 265                 270
Thr Met Pro Ala Gly Ala Thr Thr Lys Thr Glu Val Gln Glu Leu Lys
        275                 280                 285
Asp Thr Pro Ala Val Val Ser Ala Asp Ala Lys Asn Ala Leu Ile Ala
    290                 295                 300
Gly Gly Val Asp Ala Thr Asp Ala Asn Gly Ala Glu Leu Val Lys Met
305                 310                 315                 320
Ser Tyr Thr Asp Lys Asn Gly Lys Thr Ile Glu Gly Gly Tyr Ala Leu
                325                 330                 335
Lys Ala Gly Asp Lys Tyr Tyr Ala Ala Asp Tyr Asp Glu Ala Thr Gly
            340                 345                 350
Ala Ile Lys Ala Lys Thr Thr Ser Tyr Thr Ala Ala Asp Gly Thr Thr
        355                 360                 365
Lys Thr Ala Ala Asn Gln Leu Gly Gly Val Asp Gly Lys Thr Glu Val
    370                 375                 380
Val Thr Ile Asp Gly Lys Thr Tyr Asn Ala Ser Lys Ala Ala Gly His
385                 390                 395                 400
Asp Phe Lys Ala Gln Pro Glu Leu Ala Glu Ala Ala Ala Lys Thr Thr
                405                 410                 415
Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala
            420                 425                 430
Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile
        435                 440                 445
Thr Asn Leu Gly Asn Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg
    450                 455                 460
Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala
465                 470                 475                 480
Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln
                485                 490                 495
Val Pro Gln Asn Val Leu Ser Leu Leu Ala Met Glu Trp Glu Asn Ile
            500                 505                 510
Pro Ser Ile Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile
        515                 520                 525
Glu
```

<210> 343
<211> 551
<212> PRT
<213> Artificial Sequence

<220>
<223> STF2.2xH1C1

<400> 343
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
1 5 10 15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
20 25 30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
35 40 45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
50 55 60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65 70 75 80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
85 90 95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
100 105 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
115 120 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
130 135 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145 150 155 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val Gln
165 170 175
Lys Ala Tyr Asp Val Lys Asp Thr Ala Val Thr Thr Lys Ala Tyr Ala
180 185 190
Asn Asn Gly Thr Thr Leu Asp Val Ser Gly Leu Asp Asp Ala Ala Ile
195 200 205
Lys Ala Ala Thr Gly Gly Thr Asn Gly Thr Ala Ser Val Thr Gly Gly
210 215 220
Ala Val Lys Phe Asp Ala Asp Asn Asn Lys Tyr Phe Val Thr Ile Gly
225 230 235 240
Gly Phe Thr Gly Ala Asp Ala Ala Lys Asn Gly Asp Tyr Glu Val Asn
245 250 255
Val Ala Thr Asp Gly Thr Val Thr Leu Ala Ala Gly Ala Thr Lys Thr
260 265 270
Thr Met Pro Ala Gly Ala Thr Thr Lys Thr Glu Val Gln Glu Leu Lys
275 280 285
Asp Thr Pro Ala Val Val Ser Ala Asp Ala Lys Asn Ala Leu Ile Ala
290 295 300
Gly Gly Val Asp Ala Thr Asp Ala Asn Gly Ala Glu Leu Val Lys Met
305 310 315 320
Ser Tyr Thr Asp Lys Asn Gly Lys Thr Ile Glu Gly Gly Tyr Ala Leu
325 330 335
Lys Ala Gly Asp Lys Tyr Tyr Ala Ala Asp Tyr Asp Glu Ala Thr Gly
340 345 350
Ala Ile Lys Ala Lys Thr Thr Ser Tyr Thr Ala Ala Asp Gly Thr Thr
355 360 365
Lys Thr Ala Ala Asn Gln Leu Gly Gly Val Asp Gly Lys Thr Glu Val
370 375 380
Val Thr Ile Asp Gly Lys Thr Tyr Asn Ala Ser Lys Ala Ala Gly His

385 390 395 400
Asp Phe Lys Ala Gln Pro Glu Leu Ala Glu Ala Ala Ala Lys Thr Thr
405 410 415
Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala
420 425 430
Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile
435 440 445
Thr Asn Leu Gly Asn Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg
450 455 460
Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala
465 470 475 480
Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln
485 490 495
Val Pro Gln Asn Val Leu Ser Leu Leu Ala Met Glu Trp Glu Asn Ile
500 505 510
Pro Ser Ile Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile
515 520 525
Glu Glu Trp Glu Asn Ile Pro Ser Ile Gln Ser Arg Gly Leu Phe Gly
530 535 540
Ala Ile Ala Gly Phe Ile Glu
545 550

<210> 344
<211> 573
<212> PRT
<213> Artificial Sequence

<220>
<223> STF2.3xH1C1

<400> 344
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
1 5 10 15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
20 25 30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
35 40 45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
50 55 60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65 70 75 80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
85 90 95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
100 105 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
115 120 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
130 135 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145 150 155 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val Gln
165 170 175
Lys Ala Tyr Asp Val Lys Asp Thr Ala Val Thr Thr Lys Ala Tyr Ala
180 185 190
Asn Asn Gly Thr Thr Leu Asp Val Ser Gly Leu Asp Asp Ala Ala Ile
195 200 205
Lys Ala Ala Thr Gly Gly Thr Asn Gly Thr Ala Ser Val Thr Gly Gly

```
    210                 215                 220
Ala Val Lys Phe Asp Ala Asp Asn Asn Lys Tyr Phe Val Thr Ile Gly
225                 230                 235                 240
Gly Phe Thr Gly Ala Asp Ala Ala Lys Asn Gly Asp Tyr Glu Val Asn
                245                 250                 255
Val Ala Thr Asp Gly Thr Val Thr Leu Ala Ala Gly Ala Thr Lys Thr
            260                 265                 270
Thr Met Pro Ala Gly Ala Thr Thr Lys Thr Glu Val Gln Glu Leu Lys
        275                 280                 285
Asp Thr Pro Ala Val Val Ser Ala Asp Ala Lys Asn Ala Leu Ile Ala
    290                 295                 300
Gly Gly Val Asp Ala Thr Asp Ala Asn Gly Ala Glu Leu Val Lys Met
305                 310                 315                 320
Ser Tyr Thr Asp Lys Asn Gly Lys Thr Ile Glu Gly Gly Tyr Ala Leu
                325                 330                 335
Lys Ala Gly Asp Lys Tyr Tyr Ala Ala Asp Tyr Asp Glu Ala Thr Gly
            340                 345                 350
Ala Ile Lys Ala Lys Thr Thr Ser Tyr Thr Ala Ala Asp Gly Thr Thr
        355                 360                 365
Lys Thr Ala Ala Asn Gln Leu Gly Gly Val Asp Gly Lys Thr Glu Val
    370                 375                 380
Val Thr Ile Asp Gly Lys Thr Tyr Asn Ala Ser Lys Ala Ala Gly His
385                 390                 395                 400
Asp Phe Lys Ala Gln Pro Glu Leu Ala Glu Ala Ala Ala Lys Thr Thr
                405                 410                 415
Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala
            420                 425                 430
Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile
        435                 440                 445
Thr Asn Leu Gly Asn Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg
    450                 455                 460
Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala
465                 470                 475                 480
Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln
                485                 490                 495
Val Pro Gln Asn Val Leu Ser Leu Leu Ala Met Glu Trp Glu Asn Ile
            500                 505                 510
Pro Ser Ile Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile
        515                 520                 525
Glu Glu Trp Glu Asn Ile Pro Ser Ile Gln Ser Arg Gly Leu Phe Gly
    530                 535                 540
Ala Ile Ala Gly Phe Ile Glu Glu Trp Glu Asn Ile Pro Ser Ile Gln
545                 550                 555                 560
Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu
                565                 570
```

<210> 345
<211> 595
<212> PRT
<213> Artificial Sequence

<220>
<223> STF2.4xH1C1

<400> 345

```
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
 1                5                  10                  15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
```

```
            20                  25                  30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
        35                  40                  45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
    50                  55                  60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65                  70                  75                  80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
                85                  90                  95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
            100                 105                 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
        115                 120                 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
    130                 135                 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145                 150                 155                 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val Gln
                165                 170                 175
Lys Ala Tyr Asp Val Lys Asp Thr Ala Val Thr Thr Lys Ala Tyr Ala
            180                 185                 190
Asn Asn Gly Thr Thr Leu Asp Val Ser Gly Leu Asp Asp Ala Ala Ile
        195                 200                 205
Lys Ala Ala Thr Gly Gly Thr Asn Gly Thr Ala Ser Val Thr Gly Gly
    210                 215                 220
Ala Val Lys Phe Asp Ala Asp Asn Asn Lys Tyr Phe Val Thr Ile Gly
225                 230                 235                 240
Gly Phe Thr Gly Ala Asp Ala Ala Lys Asn Gly Asp Tyr Glu Val Asn
                245                 250                 255
Val Ala Thr Asp Gly Thr Val Thr Leu Ala Ala Gly Ala Thr Lys Thr
            260                 265                 270
Thr Met Pro Ala Gly Ala Thr Thr Lys Thr Glu Val Gln Glu Leu Lys
        275                 280                 285
Asp Thr Pro Ala Val Val Ser Ala Asp Ala Lys Asn Ala Leu Ile Ala
    290                 295                 300
Gly Gly Val Asp Ala Thr Asp Ala Asn Gly Ala Glu Leu Val Lys Met
305                 310                 315                 320
Ser Tyr Thr Asp Lys Asn Gly Lys Thr Ile Glu Gly Gly Tyr Ala Leu
                325                 330                 335
Lys Ala Gly Asp Lys Tyr Tyr Ala Ala Asp Tyr Asp Glu Ala Thr Gly
            340                 345                 350
Ala Ile Lys Ala Lys Thr Thr Ser Tyr Thr Ala Ala Asp Gly Thr Thr
        355                 360                 365
Lys Thr Ala Ala Asn Gln Leu Gly Gly Val Asp Gly Lys Thr Glu Val
    370                 375                 380
Val Thr Ile Asp Gly Lys Thr Tyr Asn Ala Ser Lys Ala Ala Gly His
385                 390                 395                 400
Asp Phe Lys Ala Gln Pro Glu Leu Ala Glu Ala Ala Ala Lys Thr Thr
                405                 410                 415
Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala
            420                 425                 430
Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile
        435                 440                 445
Thr Asn Leu Gly Asn Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg
    450                 455                 460
Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala
465                 470                 475                 480
Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln
                485                 490                 495
```

Val Pro Gln Asn Val Leu Ser Leu Leu Ala Met Glu Trp Glu Asn Ile
500 505 510
Pro Ser Ile Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile
515 520 525
Glu Glu Trp Glu Asn Ile Pro Ser Ile Gln Ser Arg Gly Leu Phe Gly
530 535 540
Ala Ile Ala Gly Phe Ile Glu Glu Trp Glu Asn Ile Pro Ser Ile Gln
545 550 555 560
Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Glu Trp Glu
565 570 575
Asn Ile Pro Ser Ile Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly
580 585 590
Phe Ile Glu
595

<210> 346
<211> 601
<212> PRT
<213> Artificial Sequence

<220>
<223> STF2.4xH1C2

<400> 346
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
1 5 10 15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
20 25 30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
35 40 45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
50 55 60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65 70 75 80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
85 90 95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
100 105 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
115 120 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
130 135 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145 150 155 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val Gln
165 170 175
Lys Ala Tyr Asp Val Lys Asp Thr Ala Val Thr Thr Lys Ala Tyr Ala
180 185 190
Asn Asn Gly Thr Thr Leu Asp Val Ser Gly Leu Asp Asp Ala Ala Ile
195 200 205
Lys Ala Ala Thr Gly Gly Thr Asn Gly Thr Ala Ser Val Thr Gly Gly
210 215 220
Ala Val Lys Phe Asp Ala Asp Asn Asn Lys Tyr Phe Val Thr Ile Gly
225 230 235 240
Gly Phe Thr Gly Ala Asp Ala Ala Lys Asn Gly Asp Tyr Glu Val Asn
245 250 255
Val Ala Thr Asp Gly Thr Val Thr Leu Ala Ala Gly Ala Thr Lys Thr
260 265 270

```
Thr Met Pro Ala Gly Ala Thr Thr Lys Thr Glu Val Gln Glu Leu Lys
        275                 280                 285
Asp Thr Pro Ala Val Val Ser Ala Asp Ala Lys Asn Ala Leu Ile Ala
    290                 295                 300
Gly Gly Val Asp Ala Thr Asp Ala Asn Gly Ala Glu Leu Val Lys Met
305                 310                 315                 320
Ser Tyr Thr Asp Lys Asn Gly Lys Thr Ile Glu Gly Gly Tyr Ala Leu
                325                 330                 335
Lys Ala Gly Asp Lys Tyr Tyr Ala Ala Asp Tyr Asp Glu Ala Thr Gly
            340                 345                 350
Ala Ile Lys Ala Lys Thr Thr Ser Tyr Thr Ala Ala Asp Gly Thr Thr
        355                 360                 365
Lys Thr Ala Ala Asn Gln Leu Gly Gly Val Asp Gly Lys Thr Glu Val
    370                 375                 380
Val Thr Ile Asp Gly Lys Thr Tyr Asn Ala Ser Lys Ala Ala Gly His
385                 390                 395                 400
Asp Phe Lys Ala Gln Pro Glu Leu Ala Glu Ala Ala Ala Lys Thr Thr
                405                 410                 415
Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala
            420                 425                 430
Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile
        435                 440                 445
Thr Asn Leu Gly Asn Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg
    450                 455                 460
Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala
465                 470                 475                 480
Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln
                485                 490                 495
Val Pro Gln Asn Val Leu Ser Leu Leu Ala Gly Cys Gly Ser Glu Trp
            500                 505                 510
Glu Asn Ile Pro Ser Ile Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala
        515                 520                 525
Gly Phe Ile Glu Glu Trp Glu Asn Ile Pro Ser Ile Gln Ser Arg Gly
    530                 535                 540
Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Glu Trp Glu Asn Ile Pro
545                 550                 555                 560
Ser Ile Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu
                565                 570                 575
Glu Trp Glu Asn Ile Pro Ser Ile Gln Ser Arg Gly Leu Phe Gly Ala
            580                 585                 590
Ile Ala Gly Phe Ile Glu Ser Gly Cys
        595                 600
```

<210> 347
<211> 19
<212> PRT
<213> Artificial Sequence

<220>
<223> HA maturation cleavage fragment (H5N1)

<400> 347

```
Arg Glu Arg Arg Arg Lys Lys Arg Gly Leu Phe Gly Ala Ile Ala Gly
 1               5                  10                  15
Phe Ile Glu
```

<210> 348
<211> 594
<212> PRT
<213> Artificial Sequence

<220>
<223> STF2.4xH5C1

<400> 348

```
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
 1               5                  10                  15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
            20                  25                  30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
        35                  40                  45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
    50                  55                  60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65                  70                  75                  80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
                85                  90                  95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
            100                 105                 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
        115                 120                 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
    130                 135                 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145                 150                 155                 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val Gln
                165                 170                 175
Lys Ala Tyr Asp Val Lys Asp Thr Ala Val Thr Thr Lys Ala Tyr Ala
            180                 185                 190
Asn Asn Gly Thr Thr Leu Asp Val Ser Gly Leu Asp Asp Ala Ala Ile
        195                 200                 205
Lys Ala Ala Thr Gly Gly Thr Asn Gly Thr Ala Ser Val Thr Gly Gly
    210                 215                 220
Ala Val Lys Phe Asp Ala Asp Asn Asn Lys Tyr Phe Val Thr Ile Gly
225                 230                 235                 240
Gly Phe Thr Gly Ala Asp Ala Ala Lys Asn Gly Asp Tyr Glu Val Asn
                245                 250                 255
Val Ala Thr Asp Gly Thr Val Thr Leu Ala Ala Gly Ala Thr Lys Thr
            260                 265                 270
Thr Met Pro Ala Gly Ala Thr Thr Lys Thr Glu Val Gln Glu Leu Lys
        275                 280                 285
Asp Thr Pro Ala Val Val Ser Ala Asp Ala Lys Asn Ala Leu Ile Ala
    290                 295                 300
Gly Gly Val Asp Ala Thr Asp Ala Asn Gly Ala Glu Leu Val Lys Met
305                 310                 315                 320
Ser Tyr Thr Asp Lys Asn Gly Lys Thr Ile Glu Gly Gly Tyr Ala Leu
                325                 330                 335
Lys Ala Gly Asp Lys Tyr Tyr Ala Ala Asp Tyr Asp Glu Ala Thr Gly
            340                 345                 350
Ala Ile Lys Ala Lys Thr Thr Ser Tyr Thr Ala Ala Asp Gly Thr Thr
        355                 360                 365
Lys Thr Ala Ala Asn Gln Leu Gly Gly Val Asp Gly Lys Thr Glu Val
    370                 375                 380
Val Thr Ile Asp Gly Lys Thr Tyr Asn Ala Ser Lys Ala Ala Gly His
385                 390                 395                 400
```

```
Asp Phe Lys Ala Gln Pro Glu Leu Ala Glu Ala Ala Ala Lys Thr Thr
                405                 410                 415
Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala
            420                 425                 430
Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile
        435                 440                 445
Thr Asn Leu Gly Asn Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg
    450                 455                 460
Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala
465                 470                 475                 480
Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln
                485                 490                 495
Val Pro Gln Asn Val Leu Ser Leu Leu Ala Glu Trp Glu Arg Glu Arg
            500                 505                 510
Arg Arg Lys Lys Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu
        515                 520                 525
Glu Trp Glu Arg Glu Arg Arg Arg Lys Lys Arg Gly Leu Phe Gly Ala
    530                 535                 540
Ile Ala Gly Phe Ile Glu Glu Trp Glu Arg Glu Arg Arg Arg Lys Lys
545                 550                 555                 560
Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Glu Trp Glu Arg
                565                 570                 575
Glu Arg Arg Arg Lys Lys Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe
            580                 585                 590
Ile Glu
```

<210> 349
<211> 19
<212> PRT
<213> Artificial Sequence

<220>
<223> HA maturation cleavage fragment (H3N2)

```
<400> 349
Asn Val Pro Glu Lys Gln Thr Arg Gly Ile Phe Gly Ala Ile Ala Gly
 1               5                  10                  15
Phe Ile Glu
```

<210> 350
<211> 19
<212> PRT
<213> Artificial Sequence

<220>
<223> HA maturation cleavage fragment

```
<400> 350
Asn Val Pro Gln Ile Glu Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly
 1               5                  10                  15
Phe Ile Glu
```

<210> 351

<211> 19
<212> PRT
<213> Artificial Sequence

<220>
<223> HA maturation cleavage fragment (B strain)

<400> 351
Pro Ala Lys Leu Leu Lys Glu Arg Gly Phe Phe Gly Ala Ile Ala Gly
1 5 10 15
Phe Leu Glu

<210> 352
<211> 1798
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2.4xH1C1

<400> 352
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcagaa agcgtatgat 540
gtgaaagata cagcagtaac aacgaaagct tatgccaata atggtactac actggatgta 600
tcgggtcttg atgatgcagc tattaaagcg gctacgggtg gtacgaatgg tacggcttct 660
gtaaccggtg gtgcggttaa atttgacgca gataataaca agtactttgt tactattggt 720
ggctttactg gtgctgatgc cgccaaaaat ggcgattatg aagttaacgt tgctactgac 780
ggtacagtaa cccttgcggc tggcgcaact aaaaccacaa tgcctgctgg tgcgacaact 840
aaaacagaag tacaggagtt aaaagataca ccggcagttg tttcagcaga tgctaaaaat 900
gccttaattg ctggcggcgt tgacgctacc gatgctaatg gcgctgagtt ggtcaaaatg 960
tcttataccg ataaaaatgg taagacaatt gaaggcggtt atgcgcttaa agctggcgat 1020
aagtattacg ccgcagatta cgatgaagcg acaggagcaa ttaaagctaa aactacaagt 1080
tatactgctg ctgacggcac taccaaaaca gcggctaacc aactgggtgg cgtagacggt 1140
aaaaccgaag tcgttactat cgacggtaaa acctacaatg ccagcaaagc cgctggtcat 1200
gatttcaaag cacaaccaga gctggcggaa gcagccgcta aaaccaccga aaacccgctg 1260
cagaaaattg atgccgcgct ggcgcaggtg gatgcgctgc gctctgatct gggtgcggta 1320
caaaaccgtt tcaactctgc tatcaccaac ctgggcaata ccgtaaacaa tctgtctgaa 1380
gcgcgtagcc gtatcgaaga ttccgactac gcgaccgaag tttccaacat gtctcgcgcg 1440
cagattctgc agcaggccgg tacttccgtt ctggcgcagg ctaaccaggt cccgcagaac 1500
gtgctgtctc tgttagcgat ggaatgggag aacatcccta gcatccaatc tcgcggcctg 1560
tttggcgcta tcgcgggctt tatcgaagaa tgggagaaca tcccgagcat ccaatctcgc 1620
ggtctgtttg gtgcgatcgc tggtttcatc gaggagtggg agaacattcc tagcattcaa 1680
agccgtggcc tgttcggcgc tattgcaggt tttattgaag aatgggaaaa tatcccgtct 1740
atccaatccc gcggtctgtt cggcgcgatc gcaggtttca ttgaataata agctaagc 1798

<210> 353
<211> 1738
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2.3xH1C1

<400> 353
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcagaa agcgtatgat 540
gtgaaagata cagcagtaac aacgaaagct tatgccaata atggtactac actggatgta 600
tcgggtcttg atgatgcagc tattaaagcg gctacgggtg gtacgaatgg tacggcttct 660
gtaaccggtg gtgcggttaa atttgacgca gataataaca agtactttgt tactattggt 720
ggctttactg gtgctgatgc cgccaaaaat ggcgattatg aagttaacgt tgctactgac 780
ggtacagtaa cccttgcggc tggcgcaact aaaaccacaa tgcctgctgg tgcgacaact 840
aaaacagaag tacaggagtt aaaagataca ccggcagttg tttcagcaga tgctaaaaat 900
gccttaattg ctggcggcgt tgacgctacc gatgctaatg gcgctgagtt ggtcaaaatg 960
tcttataccg ataaaaatgg taagacaatt gaaggcggtt atgcgcttaa agctggcgat 1020
aagtattacg ccgcagatta cgatgaagcg acaggagcaa ttaaagctaa aactacaagt 1080
tatactgctg ctgacggcac taccaaaaca gcggctaacc aactgggtgg cgtagacggt 1140
aaaaccgaag tcgttactat cgacggtaaa acctacaatg ccagcaaagc cgctggtcat 1200
gatttcaaag cacaaccaga gctggcggaa gcagccgcta aaaccaccga aaacccgctg 1260
cagaaaattg atgccgcgct ggcgcaggtg gatgcgctgc gctctgatct gggtgcggta 1320
caaaaccgtt tcaactctgc tatcaccaac ctgggcaata ccgtaaacaa tctgtctgaa 1380
gcgcgtagcc gtatcgaaga ttccgactac gcgaccgaag tttccaacat gtctcgcgcg 1440
cagattctgc agcaggccgg tacttccgtt ctggcgcagg ctaaccaggt cccgcagaac 1500
gtgctgagcc tgttagcgat ggaatgggaa aatatcccta gcatccaatc tcgcggtctg 1560
ttcggtgcta ttgctggctt catcgaggaa tgggagaaca tcccatctat tcagtctcgc 1620
ggcctgtttg gtgcgatcgc gggttttatt gaggaatggg aaaacattcc aagcattcag 1680
tcacgtggtc ttttcggcgc catcgctggt tttatcgaat gataagctta gcccaagg 1738

<210> 354
<211> 1672
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2.2xH1C1

<400> 354
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcagaa agcgtatgat 540
gtgaaagata cagcagtaac aacgaaagct tatgccaata atggtactac actggatgta 600
tcgggtcttg atgatgcagc tattaaagcg gctacgggtg gtacgaatgg tacggcttct 660
gtaaccggtg gtgcggttaa atttgacgca gataataaca agtactttgt tactattggt 720
ggctttactg gtgctgatgc cgccaaaaat ggcgattatg aagttaacgt tgctactgac 780
ggtacagtaa cccttgcggc tggcgcaact aaaaccacaa tgcctgctgg tgcgacaact 840
aaaacagaag tacaggagtt aaaagataca ccggcagttg tttcagcaga tgctaaaaat 900
gccttaattg ctggcggcgt tgacgctacc gatgctaatg gcgctgagtt ggtcaaaatg 960

tcttataccg ataaaaatgg taagacaatt gaaggcggtt atgcgcttaa agctggcgat 1020
aagtattacg ccgcagatta cgatgaagcg acaggagcaa ttaaagctaa aactacaagt 1080
tatactgctg ctgacggcac taccaaaaca gcggctaacc aactgggtgg cgtagacggt 1140
aaaaccgaag tcgttactat cgacggtaaa acctacaatg ccagcaaagc cgctggtcat 1200
gatttcaaag cacaaccaga gctggcggaa gcagccgcta aaaccaccga aaacccgctg 1260
cagaaaattg atgccgcgct ggcgcaggtg gatgcgctgc gctctgatct gggtgcggta 1320
caaaaccgtt tcaactctgc tatcaccaac ctgggcaata ccgtaaacaa tctgtctgaa 1380
gcgcgtagcc gtatcgaaga ttccgactac gcgaccgaag tttccaacat gtctcgcgcg 1440
cagattctgc agcaggccgg tacttccgtt ctggcgcagg ctaaccaggt cccgcagaac 1500
gtgctgagcc tgttagcgat ggaatgggaa aatatcccta gcatccaatc tcgcggtctg 1560
ttcggtgcta ttgctggctt catcgaggaa tgggagaaca tcccatctat tcagtctcgc 1620
ggcctgtttg gtgcgatcgc gggttttatt gagtgataag cttagcccaa gg 1672

<210> 355
<211> 1606
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2.1xH1C1

<400> 355
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcagaa agcgtatgat 540
gtgaaagata cagcagtaac aacgaaagct tatgccaata atggtactac actggatgta 600
tcgggtcttg atgatgcagc tattaaagcg gctacgggtg gtacgaatgg tacggcttct 660
gtaaccggtg gtgcggttaa atttgacgca gataataaca agtactttgt tactattggt 720
ggctttactg gtgctgatgc cgccaaaaat ggcgattatg aagttaacgt tgctactgac 780
ggtacagtaa cccttgcggc tggcgcaact aaaaccacaa tgcctgctgg tgcgacaact 840
aaaacagaag tacaggagtt aaaagataca ccggcagttg tttcagcaga tgctaaaaat 900
gccttaattg ctggcggcgt tgacgctacc gatgctaatg gcgctgagtt ggtcaaaatg 960
tcttataccg ataaaaatgg taagacaatt gaaggcggtt atgcgcttaa agctggcgat 1020
aagtattacg ccgcagatta cgatgaagcg acaggagcaa ttaaagctaa aactacaagt 1080
tatactgctg ctgacggcac taccaaaaca gcggctaacc aactgggtgg cgtagacggt 1140
aaaaccgaag tcgttactat cgacggtaaa acctacaatg ccagcaaagc cgctggtcat 1200
gatttcaaag cacaaccaga gctggcggaa gcagccgcta aaaccaccga aaacccgctg 1260
cagaaaattg atgccgcgct ggcgcaggtg gatgcgctgc gctctgatct gggtgcggta 1320
caaaaccgtt tcaactctgc tatcaccaac ctgggcaata ccgtaaacaa tctgtctgaa 1380
gcgcgtagcc gtatcgaaga ttccgactac gcgaccgaag tttccaacat gtctcgcgcg 1440
cagattctgc agcaggccgg tacttccgtt ctggcgcagg ctaaccaggt cccgcagaac 1500
gtgctgagcc tgttagcgat ggaatgggaa aatatcccta gcatccaatc tcgcggtctg 1560
ttcggtgcta ttgctggctt catcgagtga taagcttagc ccaagg 1606

<210> 356
<211> 1816
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2.4xH1C1

<400> 356

atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcagaa agcgtatgat 540
gtgaaagata cagcagtaac aacgaaagct tatgccaata atggtactac actggatgta 600
tcgggtcttg atgatgcagc tattaaagcg gctacgggtg gtacgaatgg tacggcttct 660
gtaaccggtg gtgcggttaa atttgacgca gataataaca agtactttgt tactattggt 720
ggctttactg gtgctgatgc cgccaaaaat ggcgattatg aagttaacgt tgctactgac 780
ggtacagtaa cccttgcggc tggcgcaact aaaaccacaa tgcctgctgg tgcgacaact 840
aaaacagaag tacaggagtt aaaagataca ccggcagttg tttcagcaga tgctaaaaat 900
gccttaattg ctggcggcgt tgacgctacc gatgctaatg gcgctgagtt ggtcaaaatg 960
tcttataccg ataaaaatgg taagacaatt gaaggcggtt atgcgcttaa agctggcgat 1020
aagtattacg ccgcagatta cgatgaagcg acaggagcaa ttaaagctaa aactacaagt 1080
tatactgctg ctgacggcac taccaaaaca gcggctaacc aactgggtgg cgtagacggt 1140
aaaaccgaag tcgttactat cgacggtaaa acctacaatg ccagcaaagc cgctggtcat 1200
gatttcaaag cacaaccaga gctggcggaa gcagccgcta aaaccaccga aaacccgctg 1260
cagaaaattg atgccgcgct ggcgcaggtg gatgcgctgc gctctgatct gggtgcggta 1320
caaaaccgtt tcaactctgc tatcaccaac ctgggcaata ccgtaaacaa tctgtctgaa 1380
gcgcgtagcc gtatcgaaga ttccgactac gcgaccgaag tttccaacat gtctcgcgcg 1440
cagattctgc agcaggccgg tacttccgtt ctggcgcagg ctaaccaggt cccgcagaac 1500
gtgctgtctc tgttagcggg ttgtggttcc gagtgggaaa atattccgtc tatccagagc 1560
cgtggtctgt tcggcgcaat tgctggcttc attgaagaat gggaaaacat cccgtccatc 1620
cagagccgtg gcctgttcgg cgccattgct ggtttcatcg aggaatggga aaacattccg 1680
tccatccagt cccgcggtct gtttggcgct atcgccggtt tcattgagga atgggaaaat 1740
atcccttcca tccagtctcg tggtctgttc ggcgcgattg caggctttat cgaatctggt 1800
tgctaataag ctaagc 1816

<210> 357
<211> 78
<212> PRT
<213> Bacterial lipoprotein of E. coli

<400> 357
Met Lys Ala Thr Lys Leu Val Leu Gly Ala Val Ile Leu Gly Ser Thr
1 5 10 15
Leu Leu Ala Gly Cys Ser Ser Asn Ala Lys Ile Asp Gln Leu Ser Ser
20 25 30
Asp Val Gln Thr Leu Asn Ala Lys Val Asp Gln Leu Ser Asn Asp Val
35 40 45
Asn Ala Met Arg Ser Asp Val Gln Ala Ala Lys Asp Asp Ala Ala Arg
50 55 60
Ala Asn Gln Arg Leu Asp Asn Met Ala Thr Lys Tyr Arg Lys
65 70 75

<210> 358
<211> 26
<212> PRT
<213> Artificial Sequence

<220>
<223> H1C1 native peptide

<400> 358

Ser Leu Trp Ser Glu Glu Asn Ile Pro Ser Ile Gln Ser Arg Gly Leu
1 5 10 15
Phe Gly Ala Ile Ala Gly Phe Ile Glu Glu
20 25

<210> 359
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 359
Gly Gly Lys Ser Gly Arg Thr Gly
1 5

<210> 360
<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 360
Lys Gly Tyr Asp Trp Leu Val Val Gly
1 5

<210> 361
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 361
Glu Asp Met Val Tyr Arg Ile Gly Val Pro
1 5 10

<210> 362
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 362
Val Lys Leu Ser Gly Ser
1 5

<210> 363

<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 363
Gly Met Leu Ser Leu Ala Leu Phe
1 5

<210> 364
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 364
Cys Val Val Gly Ser Val Arg
1 5

<210> 365
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 365
Ile Val Arg Gly Cys Leu Gly Trp
1 5

<210> 366
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 366
Ala Ala Glu Glu Arg Thr Leu Gly
1 5

<210> 367
<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 367
Trp Ala Arg Val Val Gly Trp Leu Arg
1 5

<210> 368
<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 368
Ser Glu Gly Tyr Arg Leu Phe Gly Gly
1 5

<210> 369
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 369
Leu Val Gly Gly Val Val Arg Arg Gly Ser
1 5 10

<210> 370
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 370
Gly Arg Val Asn Asp Leu Trp Leu Ala Ala
1 5 10

<210> 371
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 371
Ser Gly Trp Met Leu Trp Arg Glu Gly Ser
1 5 10

<210> 372
<211> 10

<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 372
Glu Arg Met Glu Asp Arg Gly Gly Asp Leu
1 5 10

<210> 373
<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 373
Lys Leu Cys Cys Phe Thr Glu Cys Met
1 5

<210> 374
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 374
Ala Val Gly Ser Met Glu Arg Gly Arg Gly
1 5 10

<210> 375
<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 375
Arg Asp Trp Val Gly Gly Asp Leu Val
1 5

<210> 376
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 376

Phe Phe Glu Val Ala Lys Ile Ser Gln Gln
1 5 10

<210> 377
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 377
Trp Trp Tyr Trp Cys
1 5

<210> 378
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 378
Met His Leu Cys Ser His Ala
1 5

<210> 379
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 379
Trp Leu Phe Arg Arg Ile Gly
1 5

<210> 380
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 380
Tyr Trp Phe Trp Arg Ile Gly
1 5

<210> 381
<211> 7
<212> PRT

<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 381
Met His Leu Tyr Cys Ile Ala
1 5

<210> 382
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 382
Trp Pro Leu Phe Pro Trp Ile Val
1 5

<210> 383
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 383
Asp Met Arg Ser His Ala Arg
1 5

<210> 384
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 384
Met His Leu Cys Thr His Ala
1 5

<210> 385
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 385
Asn Leu Phe Pro Phe Tyr

1 5

<210> 386
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 386
Met His Leu Cys Thr Arg Ala
1 5

<210> 387
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 387
Arg His Leu Trp Tyr His Ala
1 5

<210> 388
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 388
Trp Pro Phe Ser Ala Tyr Trp
1 5

<210> 389
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 389
Trp Tyr Leu Arg Gly Ser
1 5

<210> 390
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 390
Gly Lys Gly Thr Asp Leu Gly
1 5

<210> 391
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 391
Ile Phe Val Arg Met Arg
1 5

<210> 392
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 392
Trp Leu Phe Arg Pro Val Phe
1 5

<210> 393
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 393
Phe Leu Gly Trp Leu Met Gly
1 5

<210> 394
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 394
Met His Leu Trp His His Ala
1 5

<210> 395
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 395
Trp Trp Phe Pro Trp Lys Ala
1 5

<210> 396
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 396
Trp Tyr Leu Pro Trp Leu Gly
1 5

<210> 397
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 397
Trp Pro Phe Pro Arg Thr Phe
1 5

<210> 398
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 398
Trp Pro Phe Pro Ala Tyr Trp
1 5

<210> 399
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 399
Phe Leu Gly Leu Arg Trp Leu
1 5

<210> 400
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 400
Ser Arg Thr Asp Val Gly Val Leu Glu Val
1 5 10

<210> 401
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 401
Arg Glu Lys Val Ser Arg Gly Asp Lys Gly
1 5 10

<210> 402
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 402
Asp Trp Asp Ala Val Glu Ser Glu Tyr Met
1 5 10

<210> 403
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 403
Val Ser Ser Ala Gln Glu Val Arg Val Pro
1 5 10

<210> 404
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 404
Leu Thr Tyr Gly Gly Leu Glu Ala Leu Gly
1 5 10

<210> 405
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 405
Val Glu Glu Tyr Ser Ser Ser Gly Val Ser
1 5 10

<210> 406
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 406
Val Cys Glu Val Ser Asp Ser Val Met Ala
1 5 10

<210> 407
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 407
Gly Gly Lys Ser Gly Arg Thr Gly
1 5

<210> 408
<211> 9
<212> PRT
<213> Artificial Sequence

<220>

<223> TLR4 peptide

<400> 408
Lys Gly Tyr Asp Trp Leu Val Val Gly
1 5

<210> 409
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 409
Glu Asp Met Val Tyr Arg Ile Gly Val Pro
1 5 10

<210> 410
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 410
Val Lys Leu Ser Gly Ser
1 5

<210> 411
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 411
Gly Met Leu Ser Leu Ala Leu Phe
1 5

<210> 412
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 412
Cys Val Val Gly Ser Val Arg
1 5

<210> 413
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 413
Ile Val Arg Gly Cys Leu Gly Trp
1 5

<210> 414
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 414
Ala Ala Glu Glu Arg Thr Leu Gly
1 5

<210> 415
<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 415
Trp Ala Arg Val Val Gly Trp Leu Arg
1 5

<210> 416
<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 416
Ser Glu Gly Tyr Arg Leu Phe Gly Gly
1 5

<210> 417
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 417
Leu Val Gly Gly Val Val Arg Arg Gly Ser
1 5 10

<210> 418
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 418
Gly Arg Val Asn Asp Leu Trp Leu Ala Ala
1 5 10

<210> 419
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 419
Ser Gly Trp Met Leu Trp Arg Glu Gly Ser
1 5 10

<210> 420
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 420
Glu Arg Met Glu Asp Arg Gly Gly Asp Leu
1 5 10

<210> 421
<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 421
Lys Leu Cys Cys Phe Thr Glu Cys Met
1 5

<210> 422

<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 422
Ala Val Gly Ser Met Glu Arg Gly Arg Gly
1 5 10

<210> 423
<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 423
Arg Asp Trp Val Gly Gly Asp Leu Val
1 5

<210> 424
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 424
Phe Phe Glu Val Ala Lys Ile Ser Gln Gln
1 5 10

<210> 425
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 425
Trp Trp Tyr Trp Cys
1 5

<210> 426
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 426
Met His Leu Cys Ser His Ala
1 5

<210> 427
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 427
Trp Leu Phe Arg Arg Ile Gly
1 5

<210> 428
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 428
Tyr Trp Phe Trp Arg Ile Gly
1 5

<210> 429
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 429
Met His Leu Tyr Cys Ile Ala
1 5

<210> 430
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 430
Trp Pro Leu Phe Pro Trp Ile Val
1 5

<210> 431
<211> 7

<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 431
Asp Met Arg Ser His Ala Arg
1 5

<210> 432
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 432
Met His Leu Cys Thr His Ala
1 5

<210> 433
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 433
Asn Leu Phe Pro Phe Tyr
1 5

<210> 434
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 434
Met His Leu Cys Thr Arg Ala
1 5

<210> 435
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 435

Arg His Leu Trp Tyr His Ala
1 5

<210> 436
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 436
Trp Pro Phe Ser Ala Tyr Trp
1 5

<210> 437
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 437
Trp Tyr Leu Arg Gly Ser
1 5

<210> 438
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 438
Gly Lys Gly Thr Asp Leu Gly
1 5

<210> 439
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 439
Ile Phe Val Arg Met Arg
1 5

<210> 440
<211> 7
<212> PRT

<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 440
Trp Leu Phe Arg Pro Val Phe
1 5

<210> 441
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 441
Phe Leu Gly Trp Leu Met Gly
1 5

<210> 442
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 442
Met His Leu Trp His His Ala
1 5

<210> 443
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 443
Trp Trp Phe Pro Trp Lys Ala
1 5

<210> 444
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 444
Trp Tyr Leu Pro Trp Leu Gly

1 5

<210> 445
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 445
Trp Pro Phe Pro Arg Thr Phe
1 5

<210> 446
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 446
Trp Pro Phe Pro Ala Tyr Trp
1 5

<210> 447
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 447
Phe Leu Gly Leu Arg Trp Leu
1 5

<210> 448
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 448
Ser Arg Thr Asp Val Gly Val Leu Glu Val
1 5 10

<210> 449
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 449
Arg Glu Lys Val Ser Arg Gly Asp Lys Gly
1 5 10

<210> 450
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 450
Asp Trp Asp Ala Val Glu Ser Glu Tyr Met
1 5 10

<210> 451
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 451
Val Ser Ser Ala Gln Glu Val Arg Val Pro
1 5 10

<210> 452
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 452
Leu Thr Tyr Gly Gly Leu Glu Ala Leu Gly
1 5 10

<210> 453
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 453
Val Glu Glu Tyr Ser Ser Ser Gly Val Ser
1 5 10

<210> 454
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR4 peptide

<400> 454
Val Cys Glu Val Ser Asp Ser Val Met Ala
1 5 10

<210> 455
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 455
Asn Pro Pro Thr Thr
1 5

<210> 456
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 456
Met Arg Arg Ile Leu
1 5

<210> 457
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 457
Met Ile Ser Ser
1

<210> 458
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 458
Arg Gly Gly Ser Lys
1 5

<210> 459
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 459
Arg Gly Gly Phe
1

<210> 460
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 460
Asn Arg Thr Val Phe
1 5

<210> 461
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 461
Asn Arg Phe Gly Leu
1 5

<210> 462
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 462
Ser Arg His Gly Arg
1 5

<210> 463
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 463
Ile Met Arg His Pro
1 5

<210> 464
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 464
Glu Val Cys Ala Pro
1 5

<210> 465
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 465
Ala Cys Gly Val Tyr
1 5

<210> 466
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 466
Cys Gly Pro Lys Leu
1 5

<210> 467
<211> 5
<212> PRT
<213> Artificial Sequence

<220>

<223> TLR2 peptide

<400> 467
Ala Gly Cys Phe Ser
1 5

<210> 468
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 468
Ser Gly Gly Leu Phe
1 5

<210> 469
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 469
Ala Val Arg Leu Ser
1 5

<210> 470
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 470
Gly Gly Lys Leu Ser
1 5

<210> 471
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 471
Val Ser Glu Gly Val
1 5

<210> 472
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 472
Lys Cys Gln Ser Phe
1 5

<210> 473
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 473
Phe Cys Gly Leu Gly
1 5

<210> 474
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 474
Pro Glu Ser Gly Val
1 5

<210> 475
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 475
Asp Pro Asp Ser Gly
1 5

<210> 476
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 476
Ile Gly Arg Phe Arg
1 5

<210> 477
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 477
Met Gly Thr Leu Pro
1 5

<210> 478
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 478
Ala Asp Thr His Gln
1 5

<210> 479
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 479
His Leu Leu Pro Gly
1 5

<210> 480
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 480
Gly Pro Leu Leu His
1 5

<210> 481

<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 481
Asn Tyr Arg Arg Trp
1 5

<210> 482
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 482
Leu Arg Gln Gly Arg
1 5

<210> 483
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 483
Ile Met Trp Phe Pro
1 5

<210> 484
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 484
Arg Val Val Ala Pro
1 5

<210> 485
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 485
Ile His Val Val Pro
1 5

<210> 486
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 486
Met Phe Gly Val Pro
1 5

<210> 487
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 487
Cys Val Trp Leu Gln
1 5

<210> 488
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 488
Ile Tyr Lys Leu Ala
1 5

<210> 489
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 489
Lys Gly Trp Phe
1

<210> 490
<211> 5

<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 490
Lys Tyr Met Pro His
1 5

<210> 491
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 491
Val Gly Lys Asn Asp
1 5

<210> 492
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 492
Thr His Lys Pro Lys
1 5

<210> 493
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 493
Ser His Ile Ala Leu
1 5

<210> 494
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 494

Ala Trp Ala Gly Thr
1 5

<210> 495
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 495
Lys Gly Gly Val Gly Pro Val Arg Arg Ser Ser Arg Leu Arg Arg Thr
1 5 10 15
Thr Gln Pro Gly
20

<210> 496
<211> 23
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 496
Gly Arg Arg Gly Leu Cys Arg Gly Cys Arg Thr Arg Gly Arg Ile Lys
1 5 10 15
Gln Leu Gln Ser Ala His Lys
20

<210> 497
<211> 22
<212> PRT
<213> Artificial Sequence

<220>
<223> TLR2 peptide

<400> 497
Arg Trp Gly Tyr His Leu Arg Asp Arg Lys Tyr Lys Gly Val Arg Ser
1 5 10 15
His Lys Gly Val Pro Arg
20

<210> 498
<211> 506
<212> PRT
<213> S. typhimurium

<400> 498
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
1 5 10 15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
20 25 30

```
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
        35                  40                  45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
    50                  55                  60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65                  70                  75                  80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
                85                  90                  95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
            100                 105                 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
        115                 120                 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
    130                 135                 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145                 150                 155                 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val Gln
                165                 170                 175
Lys Ala Tyr Asp Val Lys Asp Thr Ala Val Thr Thr Lys Ala Tyr Ala
            180                 185                 190
Asn Asn Gly Thr Thr Leu Asp Val Ser Gly Leu Asp Asp Ala Ala Ile
        195                 200                 205
Lys Ala Ala Thr Gly Gly Thr Asn Gly Thr Ala Ser Val Thr Gly Gly
    210                 215                 220
Ala Val Lys Phe Asp Ala Asp Asn Asn Lys Tyr Phe Val Thr Ile Gly
225                 230                 235                 240
Gly Phe Thr Gly Ala Asp Ala Ala Lys Asn Gly Asp Tyr Glu Val Asn
                245                 250                 255
Val Ala Thr Asp Gly Thr Val Thr Leu Ala Ala Gly Ala Thr Lys Thr
            260                 265                 270
Thr Met Pro Ala Gly Ala Thr Thr Lys Thr Glu Val Gln Glu Leu Lys
        275                 280                 285
Asp Thr Pro Ala Val Val Ser Ala Asp Ala Lys Asn Ala Leu Ile Ala
    290                 295                 300
Gly Gly Val Asp Ala Thr Asp Ala Asn Gly Ala Glu Leu Val Lys Met
305                 310                 315                 320
Ser Tyr Thr Asp Lys Asn Gly Lys Thr Ile Glu Gly Gly Tyr Ala Leu
                325                 330                 335
Lys Ala Gly Asp Lys Tyr Tyr Ala Ala Asp Tyr Asp Glu Ala Thr Gly
            340                 345                 350
Ala Ile Lys Ala Lys Thr Thr Ser Tyr Thr Ala Ala Asp Gly Thr Thr
        355                 360                 365
Lys Thr Ala Ala Asn Gln Leu Gly Gly Val Asp Gly Lys Thr Glu Val
    370                 375                 380
Val Thr Ile Asp Gly Lys Thr Tyr Asn Ala Ser Lys Ala Ala Gly His
385                 390                 395                 400
Asp Phe Lys Ala Gln Pro Glu Leu Ala Glu Ala Ala Ala Lys Thr Thr
                405                 410                 415
Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala
            420                 425                 430
Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile
        435                 440                 445
Thr Asn Leu Gly Asn Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg
    450                 455                 460
Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala
465                 470                 475                 480
Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln
                485                 490                 495
Val Pro Gln Asn Val Leu Ser Leu Leu Arg
```

500 505

<210> 499
<211> 1518
<212> DNA
<213> S. typhimurium

<400> 499
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcagaa agcgtatgat 540
gtgaaagata cagcagtaac aacgaaagct tatgccaata atggtactac actggatgta 600
tcgggtcttg atgatgcagc tattaaagcg gctacgggtg gtacgaatgg tacggcttct 660
gtaaccggtg gtgcggttaa atttgacgca gataataaca agtactttgt tactattggt 720
ggctttactg gtgctgatgc cgccaaaaat ggcgattatg aagttaacgt tgctactgac 780
ggtacagtaa cccttgcggc tggcgcaact aaaaccacaa tgcctgctgg tgcgacaact 840
aaaacagaag tacaggagtt aaaagataca ccggcagttg tttcagcaga tgctaaaaat 900
gccttaattg ctggcggcgt tgacgctacc gatgctaatg gcgctgagtt ggtcaaaatg 960
tcttataccg ataaaaatgg taagacaatt gaaggcggtt atgcgcttaa agctggcgat 1020
aagtattacg ccgcagatta cgatgaagcg acaggagcaa ttaaagctaa aactacaagt 1080
tatactgctg ctgacggcac taccaaaaca gcggctaacc aactgggtgg cgtagacggt 1140
aaaaccgaag tcgttactat cgacggtaaa acctacaatg ccagcaaagc cgctggtcat 1200
gatttcaaag cacaaccaga gctggcggaa gcagccgcta aaaccaccga aaacccgctg 1260
cagaaaattg atgccgcgct ggcgcaggtg gatgcgctgc gctctgatct gggtgcggta 1320
caaaaccgtt tcaactctgc tatcaccaac ctgggcaata ccgtaaacaa tctgtctgaa 1380
gcgcgtagcc gtatcgaaga ttccgactac gcgaccgaag tttccaacat gtctcgcgcg 1440
cagattctgc agcaggccgg tacttccgtt ctggcgcagg ctaaccaggt cccgcagaac 1500
gtgctgtctc tgttacgt 1518

<210> 500
<211> 277
<212> PRT
<213> S. typhimurium

<400> 500
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
1 5 10 15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
20 25 30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
35 40 45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
50 55 60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65 70 75 80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
85 90 95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
100 105 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
115 120 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu

```
    130                 135                 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145                 150                 155                 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val His
                165                 170                 175
Gly Ala Pro Val Asp Pro Ala Ser Pro Trp Thr Glu Asn Pro Leu Gln
            180                 185                 190
Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala Leu Arg Ser Asp Leu
        195                 200                 205
Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile Thr Asn Leu Gly Asn
    210                 215                 220
Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg Ile Glu Asp Ser Asp
225                 230                 235                 240
Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala Gln Ile Leu Gln Gln
                245                 250                 255
Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln Val Pro Gln Asn Val
            260                 265                 270
Leu Ser Leu Leu Arg
        275
```

<210> 501
<211> 832
<212> DNA
<213> S. typhimurium

<400> 501
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcatgg agcgccggtg 540
gatcctgcta gcccatggac cgaaaacccg ctgcagaaaa ttgatgccgc gctggcgcag 600
gtggatgcgc tgcgctctga tctgggtgcg gtacaaaacc gtttcaactc tgctatcacc 660
aacctgggca ataccgtaaa caatctgtct gaagcgcgta gccgtatcga agattccgac 720
tacgcgaccg aagtttccaa catgtctcgc gcgcagattt tgcagcaggc cggtacttcc 780
gttctggcgc aggctaacca ggtcccgcag aacgtgctgt ctctgttacg tg 832

<210> 502
<211> 595
<212> PRT
<213> E. coli

<400> 502
```
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Ile Thr Gln Asn
 1               5                  10                  15
Asn Ile Asn Lys Asn Gln Ser Ala Leu Ser Ser Ser Ile Glu Arg Leu
            20                  25                  30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
        35                  40                  45
Ala Ile Ala Asn Arg Phe Thr Ser Asn Ile Lys Gly Leu Thr Gln Ala
    50                  55                  60
Ala Arg Asn Ala Asn Asp Gly Ile Ser Val Ala Gln Thr Thr Glu Gly
65                  70                  75                  80
Ala Leu Ser Glu Ile Asn Asn Asn Leu Gln Arg Ile Arg Glu Leu Thr
```

```
                85                  90                  95
Val Gln Ala Ser Thr Gly Thr Asn Ser Asp Ser Asp Leu Asp Ser Ile
            100                 105                 110
Gln Asp Glu Ile Lys Ser Arg Leu Asp Glu Ile Asp Arg Val Ser Gly
        115                 120                 125
Gln Thr Gln Phe Asn Gly Val Asn Val Leu Ala Lys Asp Gly Ser Met
    130                 135                 140
Lys Ile Gln Val Gly Ala Asn Asp Gly Gln Thr Ile Thr Ile Asp Leu
145                 150                 155                 160
Lys Lys Ile Asp Ser Asp Thr Leu Gly Leu Asn Gly Phe Asn Val Asn
                165                 170                 175
Gly Ser Gly Thr Ile Ala Asn Lys Ala Ala Thr Ile Ser Asp Leu Thr
            180                 185                 190
Ala Ala Lys Met Asp Ala Ala Thr Asn Thr Ile Thr Thr Thr Asn Asn
        195                 200                 205
Ala Leu Thr Ala Ser Lys Ala Leu Asp Gln Leu Lys Asp Gly Asp Thr
    210                 215                 220
Val Thr Ile Lys Ala Asp Ala Ala Gln Thr Ala Thr Val Tyr Thr Tyr
225                 230                 235                 240
Asn Ala Ser Ala Gly Asn Phe Ser Leu Ser Asn Val Ser Asn Asn Thr
                245                 250                 255
Ser Glu Lys Ala Gly Asp Val Ala Ala Ser Leu Leu Pro Pro Ala Gly
            260                 265                 270
Gln Thr Ala Ser Gly Val Tyr Lys Ala Ala Ser Gly Glu Val Asn Phe
        275                 280                 285
Asp Val Asp Ala Asn Gly Lys Ile Thr Ile Gly Gly Gln Lys Ala Tyr
    290                 295                 300
Leu Thr Ser Asp Gly Asn Leu Thr Thr Asn Asp Ala Gly Gly Ala Thr
305                 310                 315                 320
Ala Ala Thr Leu Asp Gly Leu Phe Lys Lys Ala Gly Asp Gly Gln Ser
                325                 330                 335
Ile Gly Phe Lys Lys Thr Ala Ser Val Thr Met Gly Gly Thr Thr Tyr
            340                 345                 350
Asn Phe Lys Thr Gly Ala Asp Ala Asp Ala Ala Thr Ala Asn Ala Gly
        355                 360                 365
Val Ser Phe Thr Asp Thr Ala Ser Lys Glu Thr Val Leu Asn Lys Val
    370                 375                 380
Ala Thr Ala Lys Gln Gly Lys Ala Ala Ala Ala Asp Gly Asp Thr Ser
385                 390                 395                 400
Ala Thr Ile Thr Tyr Lys Ser Gly Val Gln Thr Tyr Gln Ala Val Phe
                405                 410                 415
Ala Ala Gly Asp Gly Thr Ala Ser Ala Lys Tyr Ala Asp Lys Ala Asp
            420                 425                 430
Val Ser Asn Ala Thr Ala Thr Tyr Thr Asp Ala Asp Gly Glu Met Thr
        435                 440                 445
Thr Ile Gly Ser Tyr Thr Thr Lys Tyr Ser Ile Asp Ala Asn Asn Gly
    450                 455                 460
Lys Val Thr Val Asp Ser Gly Thr Gly Thr Gly Lys Tyr Ala Pro Lys
465                 470                 475                 480
Val Gly Ala Glu Val Tyr Val Ser Ala Asn Gly Thr Leu Thr Thr Asp
                485                 490                 495
Ala Thr Ser Glu Gly Thr Val Thr Lys Asp Pro Leu Lys Ala Leu Asp
            500                 505                 510
Glu Ala Ile Ser Ser Ile Asp Lys Phe Arg Ser Ser Leu Gly Ala Ile
        515                 520                 525
Gln Asn Arg Leu Asp Ser Ala Val Thr Asn Leu Asn Asn Thr Thr Thr
    530                 535                 540
Asn Leu Ser Glu Ala Gln Ser Arg Ile Gln Asp Ala Asp Tyr Ala Thr
545                 550                 555                 560
```

```
Glu Val Ser Asn Met Ser Lys Ala Gln Ile Ile Gln Gln Ala Gly Asn
                565                 570                 575
Ser Val Leu Ala Lys Ala Asn Gln Val Pro Gln Gln Val Leu Ser Leu
            580                 585                 590
Leu Gln Gly
        595
```

<210> 503
<211> 1788
<212> DNA
<213> E. coli

<400> 503
atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgccgcagg tcaggcgatt gctaaccgtt ttacttctaa cattaaaggc 180
ctgactcagg ctgcacgtaa cgccaacgac ggtatttccg ttgcgcagac caccgaaggc 240
gcgctgtccg aaatcaacaa caacttacag cgtatccgtg aactgacggt tcaggcttct 300
accgggacta actccgattc agatctggac tccattcagg acgaaatcaa atcccgtctg 360
gacgaaattg accgcgtatc tggccagacc cagttcaacg gcgtgaacgt actggcgaaa 420
gacggttcaa tgaaaattca ggttggtgcg aatgacggcc agactatcac gattgatctg 480
aagaaaattg actcagatac gctggggctg aatggtttta acgtgaatgg ttccggtacg 540
atagccaata aagcggcgac cattagcgac ctgacagcag cgaaaatgga tgctgcaact 600
aatactataa ctacaacaaa taatgcgctg actgcatcaa aggcgcttga tcaactgaaa 660
gatggtgaca ctgttactat caaagcagat gctgctcaaa ctgccacggt ttatacatac 720
aatgcatcag ctggtaactt ctcactcagt aatgtatcga ataatacttc agaaaaagca 780
ggtgatgtag cagctagcct tctcccgccg gctgggcaaa ctgctagtgg tgtttataaa 840
gcagcaagcg gtgaagtgaa ctttgatgtt gatgcgaatg gtaaaatcac aatcggagga 900
cagaaagcat atttaactag tgatggtaac ttaactacaa acgatgctgg tggtgcgact 960
gcggctacgc ttgatggttt attcaagaaa gctggtgatg gtcaatcaat cgggtttaag 1020
aagactgcat cagtcacgat gggggggaaca acttataact ttaaaacggg tgctgatgct 1080
gatgctgcaa ctgctaacgc aggggtatcg ttcactgata cagctagcaa agaaaccgtt 1140
ttaaataaag tggctacagc taaacaaggc aaagcagctg cagctgacgg tgatacatcc 1200
gcaacaatta cctataaatc tggcgttcag acgtatcagg ctgtatttgc cgcaggtgac 1260
ggtactgcta gcgcaaaata tgccgataaa gctgacgttt ctaatgcaac agcaacatac 1320
actgatgctg atggtgaaat gactacaatt ggttcataca ccacgaagta ttcaatcgat 1380
gctaacaacg gcaaggtaac tgttgattct ggaactggta cgggtaaata tgcgccgaaa 1440
gtaggggctg aagtatatgt tagtgctaat ggtactttaa caacagatgc aactagcgaa 1500
ggcacagtaa caaaagatcc actgaaagct ctggatgaag ctatcagctc catcgacaaa 1560
ttccgttctt ccctgggtgc tatccagaac cgtctggatt ccgcagtcac caacctgaac 1620
aacaccacta ccaacctgtc cgaagcgcag tcccgtattc aggacgccga ctatgcgacc 1680
gaagtgtcca acatgtcgaa agcgcagatc attcagcagg ccggtaactc cgtgctggca 1740
aaagccaacc aggtaccgca gcaggttctg tctctgctgc agggttag 1788

<210> 504
<211> 506
<212> PRT
<213> s. muenchen

<400> 504
```
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
 1               5                  10                  15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
            20                  25                  30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
        35                  40                  45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
    50                  55                  60
```

```
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65                  70                  75                  80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
                85                  90                  95
Val Gln Ser Ala Asn Gly Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
            100                 105                 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
        115                 120                 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
    130                 135                 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145                 150                 155                 160
Lys Glu Ile Ser Ser Lys Thr Leu Gly Leu Asp Lys Leu Asn Val Gln
                165                 170                 175
Asp Ala Tyr Thr Pro Lys Glu Thr Ala Val Thr Val Asp Lys Thr Thr
            180                 185                 190
Tyr Lys Asn Gly Thr Asp Thr Ile Thr Ala Gln Ser Asn Thr Asp Ile
        195                 200                 205
Gln Thr Ala Ile Gly Gly Gly Ala Thr Gly Val Thr Gly Ala Asp Ile
    210                 215                 220
Lys Phe Lys Asp Gly Gln Tyr Tyr Leu Asp Val Lys Gly Gly Ala Ser
225                 230                 235                 240
Ala Gly Val Tyr Lys Ala Thr Tyr Asp Glu Thr Thr Lys Lys Val Asn
                245                 250                 255
Ile Asp Thr Thr Asp Lys Thr Pro Leu Ala Thr Ala Glu Ala Thr Ala
            260                 265                 270
Ile Arg Gly Thr Ala Thr Ile Thr His Asn Gln Ile Ala Glu Val Thr
        275                 280                 285
Lys Glu Gly Val Asp Thr Thr Thr Val Ala Ala Gln Leu Ala Ala Ala
    290                 295                 300
Gly Val Thr Gly Ala Asp Lys Asp Asn Thr Ser Leu Val Lys Leu Ser
305                 310                 315                 320
Phe Glu Asp Lys Asn Gly Lys Val Ile Asp Gly Gly Tyr Ala Val Lys
                325                 330                 335
Met Gly Asp Asp Phe Tyr Ala Ala Thr Tyr Asp Glu Lys Thr Gly Thr
            340                 345                 350
Ile Thr Ala Lys Thr Thr Thr Tyr Thr Asp Gly Ala Gly Val Ala Gln
        355                 360                 365
Thr Gly Ala Val Lys Phe Gly Gly Ala Asn Gly Lys Ser Glu Val Val
    370                 375                 380
Thr Ala Thr Asp Gly Lys Thr Tyr Leu Ala Ser Asp Leu Asp Lys His
385                 390                 395                 400
Asn Phe Arg Thr Gly Gly Glu Leu Lys Glu Val Asn Thr Asp Lys Thr
                405                 410                 415
Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Thr
            420                 425                 430
Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile
        435                 440                 445
Thr Asn Leu Gly Asn Thr Val Asn Asn Leu Ser Ser Ala Arg Ser Arg
    450                 455                 460
Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala
465                 470                 475                 480
Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln
                485                 490                 495
Val Pro Gln Asn Val Leu Ser Leu Leu Arg
            500                 505
```

<210> 505

<211> 1522
<212> DNA
<213> s. muenchen

<400> 505
aatggcacaa gtcattaata caaacagcct gtcgctgttg acccagaata acctgaacaa 60
atcccagtcc gctctgggca ccgctatcga gcgtctgtct tccggtctgc gtatcaacag 120
cgcgaaagac gatgcggcag gtcaggcgat tgctaaccgt ttcaccgcga acatcaaagg 180
tctgactcag gcttcccgta acgctaacga cggtatctcc attgcgcaga ccactgaagg 240
cgcgctgaac gaaatcaaca acaacctgca gcgtgtgcgt gaactggcgg ttcagtctgc 300
taacggtact aactcccagt ctgaccttga ctctatccag gctgaaatca cccagcgtct 360
gaacgaaatc gaccgtgtat ccggtcagac tcagttcaac ggcgtgaaag tcctggcgca 420
ggacaacacc ctgaccatcc aggttggtgc caacgacggt gaaactattg atattgattt 480
aaaagaaatt agctctaaaa cactgggact tgataagctt aatgtccagg atgcctacac 540
cccgaaagaa actgctgtaa ccgttgataa aactacctat aaaaatggta cagatactat 600
tacagcccag agcaatactg atatccaaac tgcaattggc ggtggtgcaa cgggggttac 660
tggggctgat atcaaattta aagatggtca atactattta gatgttaaag gcggtgcttc 720
tgctggtgtt tataaagcca cttatgatga aactacaaag aaagttaata ttgatacgac 780
tgataaaact ccgttagcaa ctgcggaagc tacagctatt cggggaacgg ccactataac 840
ccacaaccaa attgctgaag taacaaaaga gggtgttgat acgaccacag ttgcggctca 900
acttgctgct gcaggggtta ctggtgccga taaggacaat actagccttg taaaactatc 960
gtttgaggat aaaaacggta aggttattga tggtggctat gcagtgaaaa tgggcgacga 1020
tttctatgcc gctacatatg atgagaaaac aggtacaatt actgctaaaa caaccactta 1080
tacagatggt gctggcgttg ctcaaactgg agctgtgaaa tttggtggcg caaatggtaa 1140
atctgaagtt gttactgcta ccgatggtaa aacttactta gcaagcgacc ttgacaaaca 1200
taacttcaga acaggcggtg agcttaaaga ggttaataca gataagactg aaaacccact 1260
gcagaaaatt gatgctgcct tggcacaggt tgatacactt cgttctgacc tgggtgcggt 1320
acagaaccgt ttcaactccg ctatcaccaa cctgggcaat accgtaaata acctgtcttc 1380
tgcccgtagc cgtatcgaag attccgacta cgcgaccgaa gtctccaaca tgtctcgcgc 1440
gcagattctg cagcaggccg gtacctccgt tctggcgcag gctaaccagg ttccgcaaaa 1500
cgtcctctct ttactgcgtt aa 1522

<210> 506
<211> 588
<212> PRT
<213> Artificial Sequence

<220>
<223> pMT/STF2

<400> 506
Met Lys Leu Cys Ile Leu Leu Ala Val Val Ala Phe Val Gly Leu Ser
1 5 10 15
Leu Gly Arg Ser Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu
20 25 30
Thr Gln Asn Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile
35 40 45
Glu Arg Leu Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala
50 55 60
Ala Gly Gln Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu
65 70 75 80
Thr Gln Ala Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr
85 90 95
Thr Glu Gly Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg
100 105 110
Glu Leu Ala Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu
115 120 125
Asp Ser Ile Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg
130 135 140

```
Val Ser Gly Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp
145                 150                 155                 160
Asn Thr Leu Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp
                165                 170                 175
Ile Asp Leu Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu
            180                 185                 190
Asn Val Gln Lys Ala Tyr Asp Val Lys Asp Thr Ala Val Thr Thr Lys
        195                 200                 205
Ala Tyr Ala Asn Asn Gly Thr Thr Leu Asp Val Ser Gly Leu Asp Asp
    210                 215                 220
Ala Ala Ile Lys Ala Ala Thr Gly Gly Thr Asn Gly Thr Ala Ser Val
225                 230                 235                 240
Thr Gly Gly Ala Val Lys Phe Asp Ala Asp Asn Asn Lys Tyr Phe Val
                245                 250                 255
Thr Ile Gly Gly Phe Thr Gly Ala Asp Ala Ala Lys Asn Gly Asp Tyr
            260                 265                 270
Glu Val Asn Val Ala Thr Asp Gly Thr Val Thr Leu Ala Ala Gly Ala
        275                 280                 285
Thr Lys Thr Thr Met Pro Ala Gly Ala Thr Thr Lys Thr Glu Val Gln
    290                 295                 300
Glu Leu Lys Asp Thr Pro Ala Val Val Ser Ala Asp Ala Lys Asn Ala
305                 310                 315                 320
Leu Ile Ala Gly Gly Val Asp Ala Thr Asp Ala Asn Gly Ala Glu Leu
                325                 330                 335
Val Lys Met Ser Tyr Thr Asp Lys Asn Gly Lys Thr Ile Glu Gly Gly
            340                 345                 350
Tyr Ala Leu Lys Ala Gly Asp Lys Tyr Tyr Ala Ala Asp Tyr Asp Glu
        355                 360                 365
Ala Thr Gly Ala Ile Lys Ala Lys Thr Thr Ser Tyr Thr Ala Ala Asp
    370                 375                 380
Gly Thr Thr Lys Thr Ala Ala Asn Gln Leu Gly Gly Val Asp Gly Lys
385                 390                 395                 400
Thr Glu Val Val Thr Ile Asp Gly Lys Thr Tyr Asn Ala Ser Lys Ala
                405                 410                 415
Ala Gly His Asp Phe Lys Ala Gln Pro Glu Leu Ala Glu Ala Ala Ala
            420                 425                 430
Lys Thr Thr Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala Gln
        435                 440                 445
Val Asp Ala Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe Asn
    450                 455                 460
Ser Ala Ile Thr Asn Leu Gly Asn Thr Val Asn Asn Leu Ser Glu Ala
465                 470                 475                 480
Arg Ser Arg Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn Met
                485                 490                 495
Ser Arg Ala Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala Gln
            500                 505                 510
Ala Asn Gln Val Pro Gln Asn Val Leu Ser Leu Leu Arg Lys Gly Asn
        515                 520                 525
Ser Lys Leu Glu Gly Gln Leu Glu Phe Pro Arg Thr Ser Pro Val Trp
    530                 535                 540
Trp Asn Ser Ala Asp Ile Gln His Ser Gly Gly Arg Ser Ser Leu Glu
545                 550                 555                 560
Gly Pro Arg Phe Glu Gly Lys Pro Ile Pro Asn Pro Leu Leu Gly Leu
                565                 570                 575
Asp Ser Thr Arg Thr Gly His His His His His His
            580                 585
```

<210> 507

<211> 1764
<212> DNA
<213> Artificial Sequence

<220>
<223> pMT/STF2

<400> 507
atgaagttat gcatattact ggccgtcgtg gcctttgttg gcctctcgct cgggagatct 60
gcacaagtaa tcaacactaa cagtctgtcg ctgctgaccc agaataacct gaacaaatcc 120
cagtccgcac tgggcaccgc tatcgagcgt ctgtcttctg gtctgcgtat caacagcgcg 180
aaagacgatg cggcaggtca ggcgattgct aaccgtttca ccgcgaacat caaaggtctg 240
actcaggctt cccgtaacgc taacgacggt atctccattg cgcagaccac tgaaggcgcg 300
ctgaacgaaa tcaacaacaa cctgcagcgt gtgcgtgaac tggcggttca gtctgctaac 360
agcaccaact cccagtctga cctcgactcc atccaggctg aaatcaccca gcgcctgaac 420
gaaatcgacc gtgtatccgg ccagactcag ttcaacggcg tgaaagtcct ggcgcaggac 480
aacaccctga ccatccaggt tggcgccaac gacggtgaaa ctatcgatat cgatctgaag 540
cagatcaact ctcagaccct gggtctggac tcactgaacg tgcagaaagc gtatgatgtg 600
aaagatacag cagtaacaac gaaagcttat gccaataatg gtactacact ggatgtatcg 660
ggtcttgatg atgcagctat taaagcggct acgggtggta cgaatggtac ggcttctgta 720
accggtggtg cggttaaatt tgacgcagat aataacaagt actttgttac tattggtggc 780
tttactggtg ctgatgccgc caaaaatggc gattatgaag ttaacgttgc tactgacggt 840
acagtaaccc ttgcggctgg cgcaactaaa accacaatgc ctgctggtgc gacaactaaa 900
acagaagtac aggagttaaa agatacaccg gcagttgttt cagcagatgc taaaaatgcc 960
ttaattgctg gcggcgttga cgctaccgat gctaatggcg ctgagttggt caaaatgtct 1020
tataccgata aaaatggtaa gacaattgaa ggcggttatg cgcttaaagc tggcgataag 1080
tattacgccg cagattacga tgaagcgaca ggagcaatta aagctaaaac tacaagttat 1140
actgctgctg acggcactac caaaacagcg gctaaccaac tgggtggcgt agacggtaaa 1200
accgaagtcg ttactatcga cggtaaaacc tacaatgcca gcaaagccgc tggtcatgat 1260
ttcaaagcac aaccagagct ggcggaagca gccgctaaaa ccaccgaaaa cccgctgcag 1320
aaaattgatg ccgcgctggc gcaggtggat gcgctgcgct ctgatctggg tgcggtacaa 1380
aaccgtttca actctgctat caccaacctg ggcaataccg taaacaatct gtctgaagcg 1440
cgtagccgta tcgaagattc cgactacgcg accgaagttt ccaacatgtc tcgcgcgcag 1500
attctgcagc aggccggtac ttccgttctg gcgcaggcta accaggtccc gcagaacgtg 1560
ctgtctctgt tacgtaaggg caattcgaag cttgaaggtc aattggaatt ccctaggact 1620
agtccagtgt ggtggaattc tgcagatatc cagcacagtg gcggccgctc gagtctagag 1680
ggcccgcggt tcgaaggtaa gcctatccct aaccctctcc tcggtctcga ttctacgcgt 1740
accggtcatc atcaccatca ccat 1764

<210> 508
<211> 97
<212> PRT
<213> Artificial Sequence

<220>
<223> M2 Influenza A

<400> 508
Met Ser Leu Leu Thr Glu Val Glu Thr Pro Ile Arg Asn Glu Trp Gly
1 5 10 15
Cys Arg Cys Asn Asp Ser Ser Asp Pro Leu Val Val Ala Ala Ser Ile
20 25 30
Ile Gly Ile Leu His Leu Ile Leu Trp Ile Leu Asp Arg Leu Phe Phe
35 40 45
Lys Cys Ile Tyr Arg Leu Phe Lys His Gly Leu Lys Arg Gly Pro Ser
50 55 60
Thr Glu Gly Val Pro Glu Ser Met Arg Glu Glu Tyr Arg Lys Glu Gln
65 70 75 80
Gln Asn Ala Val Asp Ala Asp Asp Ser His Phe Val Ser Ile Glu Leu

85 90 95
Glu

<210> 509
<211> 294
<212> DNA
<213> Artificial Sequence

<220>
<223> M2 Influenza A

<400> 509
atgagccttc taaccgaggt cgaaacacct atcagaaacg aatgggggtg cagatgcaac 60
gattcaagtg acccgcttgt tgttgccgcg agtatcattg ggatcttgca cttgatattg 120
tggattcttg atcgtctttt tttcaaatgc atctatcgac tcttcaaaca cggccttaaa 180
agagggcctt ctacggaagg agtacctgag tctatgaggg aagaatatcg aaaggaacag 240
cagaatgctg tggatgctga cgacagtcat tttgtcagca tagagttgga gtaa 294

<210> 510
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> M2

<400> 510
Ser Leu Leu Thr Glu Val Glu Thr Pro Ile Arg Asn Glu Trp Gly Cys
1 5 10 15
Arg Cys Asn Asp Ser Ser Asp Pro
20

<210> 511
<211> 75
<212> DNA
<213> Artificial Sequence

<220>
<223> M2

<400> 511
atgagcctgc tgaccgaggt cgaaacaccg atccgcaacg aatgggggtg ccgctgcaac 60
gattcaagtg acccg 75

<210> 512
<211> 11
<212> PRT
<213> Artificial Sequence

<220>
<223> M2

<400> 512
Glu Val Glu Thr Pro Ile Arg Asn Glu Trp Gly
1 5 10

<210> 513
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> M2

<400> 513
gaggtcgaaa cacctatcag aaacgaatgg 30

<210> 514
<211> 315
<212> DNA
<213> Artificial Sequence

<220>
<223> 4XM2e

<400> 514
agcttgctga ctgaggttga gaccccgatt cgcaacgaat ggggttcccg ttccaacgat 60
tcttccgacc cgctcgagag cctgttgacc gaggttgaaa cccctatccg taatgaatgg 120
ggctcccgta gcaacgactc ttctgaccca ggatcctccc tcttgaccga agtggaaacg 180
cctattcgta atgagtgggg ttctcgtagc aatgacagct cggacccgga gctctcgctg 240
ctgacggaag tggagactcc gatccgtaac gagtggggct ctcgctctaa cgatagctca 300
gacccgtcta gataa 315

<210> 515
<211> 104
<212> PRT
<213> Artificial Sequence

<220>
<223> 4XM2e

<400> 515
Ser Leu Leu Thr Glu Val Glu Thr Pro Ile Arg Asn Glu Trp Gly Ser
1 5 10 15
Arg Ser Asn Asp Ser Ser Asp Pro Leu Glu Ser Leu Leu Thr Glu Val
20 25 30
Glu Thr Pro Ile Arg Asn Glu Trp Gly Ser Arg Ser Asn Asp Ser Ser
35 40 45
Asp Pro Gly Ser Ser Leu Leu Thr Glu Val Glu Thr Pro Ile Arg Asn
50 55 60
Glu Trp Gly Ser Arg Ser Asn Asp Ser Ser Asp Pro Glu Leu Ser Leu
65 70 75 80
Leu Thr Glu Val Glu Thr Pro Ile Arg Asn Glu Trp Gly Ser Arg Ser
85 90 95
Asn Asp Ser Ser Asp Pro Ser Arg
100

<210> 516
<211> 9
<212> PRT
<213> Artificial Sequence

<220>

<223> M2

<400> 516
Glu Val Glu Thr Pro Ile Arg Asn Glu
1 5

<210> 517
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> M2

<400> 517
Glu Val Glu Thr Pro Ile Arg Asn
1 5

<210> 518
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> M2

<400> 518
Met Lys Ala Thr Lys Leu Val Leu Gly Ala Val Ile Leu Gly Ser Thr
1 5 10 15
Leu Leu Ala Gly Cys Ser Ser Asn
20

<210> 519
<211> 75
<212> DNA
<213> Artificial Sequence

<220>
<223> M2

<400> 519
atgaaagcta ctaaactggt actgggcgcg gtaatcctgg gttctactct gctgctggca 60
ggttgctcca gcaac 75

<210> 520
<211> 48
<212> DNA
<213> Artificial Sequence

<220>
<223> Linker

<400> 520
aagggcaatt cgaagcttga aggtcaattg gaattcccta ggactagt 48

<210> 521
<211> 16
<212> PRT
<213> Artificial Sequence

<220>
<223> Linker

<400> 521
Lys Gly Asn Ser Lys Leu Glu Gly Gln Leu Glu Phe Pro Arg Thr Ser
1 5 10 15

<210> 522
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Linker

<400> 522
gaattctgca gatatccagc acagtggcgg ccgctc 36

<210> 523
<211> 12
<212> PRT
<213> Artificial Sequence

<220>
<223> Linker

<400> 523
Glu Phe Cys Arg Tyr Pro Ala Gln Trp Arg Pro Leu
1 5 10

<210> 524
<211> 12
<212> PRT
<213> Artificial Sequence

<220>
<223> Linker

<400> 524
Glu Phe Ser Arg Tyr Pro Ala Gln Trp Arg Pro Leu
1 5 10

<210> 525

<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Linker

<400> 525
gaattctcta gatatccagc acagtggcgg ccgctc 36

<210> 526
<211> 45
<212> PRT
<213> Artificial Sequence

<220>
<223> Linker

<400> 526
Lys Gly Asn Ser Lys Leu Glu Gly Gln Leu Glu Phe Pro Arg Thr Ser
1 5 10 15
Pro Val Trp Trp Asn Ser Ala Asp Ile Gln His Ser Gly Gly Arg Gln
20 25 30
Cys Asp Gly Tyr Leu Gln Asn Ser Pro Leu Arg Pro Leu
35 40 45

<210> 527
<211> 135
<212> DNA
<213> Artificial Sequence

<220>
<223> Linker

<400> 527
aagggcaatt cgaagcttga aggtcaattg gaattcccta ggactagtcc agtgtggtgg 60
aattctgcag atatccagca cagtggcggc cgccagtgtg atggatatct gcagaattcg 120
cccttgcggc cgctc 135

<210> 528
<211> 611
<212> PRT
<213> Artificial Sequence

<220>
<223> STF2-4XM2e

<400> 528
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
1 5 10 15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
20 25 30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
35 40 45

```
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
    50                  55                  60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65                  70                  75                  80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
                85                  90                  95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
            100                 105                 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
        115                 120                 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
    130                 135                 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145                 150                 155                 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val Gln
                165                 170                 175
Lys Ala Tyr Asp Val Lys Asp Thr Ala Val Thr Thr Lys Ala Tyr Ala
            180                 185                 190
Asn Asn Gly Thr Thr Leu Asp Val Ser Gly Leu Asp Asp Ala Ala Ile
        195                 200                 205
Lys Ala Ala Thr Gly Gly Thr Asn Gly Thr Ala Ser Val Thr Gly Gly
    210                 215                 220
Ala Val Lys Phe Asp Ala Asp Asn Asn Lys Tyr Phe Val Thr Ile Gly
225                 230                 235                 240
Gly Phe Thr Gly Ala Asp Ala Ala Lys Asn Gly Asp Tyr Glu Val Asn
                245                 250                 255
Val Ala Thr Asp Gly Thr Val Thr Leu Ala Ala Gly Ala Thr Lys Thr
            260                 265                 270
Thr Met Pro Ala Gly Ala Thr Thr Lys Thr Glu Val Gln Glu Leu Lys
        275                 280                 285
Asp Thr Pro Ala Val Val Ser Ala Asp Ala Lys Asn Ala Leu Ile Ala
    290                 295                 300
Gly Gly Val Asp Ala Thr Asp Ala Asn Gly Ala Glu Leu Val Lys Met
305                 310                 315                 320
Ser Tyr Thr Asp Lys Asn Gly Lys Thr Ile Glu Gly Gly Tyr Ala Leu
                325                 330                 335
Lys Ala Gly Asp Lys Tyr Tyr Ala Ala Asp Tyr Asp Glu Ala Thr Gly
            340                 345                 350
Ala Ile Lys Ala Lys Thr Thr Ser Tyr Thr Ala Ala Asp Gly Thr Thr
        355                 360                 365
Lys Thr Ala Ala Asn Gln Leu Gly Gly Val Asp Gly Lys Thr Glu Val
    370                 375                 380
Val Thr Ile Asp Gly Lys Thr Tyr Asn Ala Ser Lys Ala Ala Gly His
385                 390                 395                 400
Asp Phe Lys Ala Gln Pro Glu Leu Ala Glu Ala Ala Ala Lys Thr Thr
                405                 410                 415
Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala
            420                 425                 430
Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile
        435                 440                 445
Thr Asn Leu Gly Asn Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg
    450                 455                 460
Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala
465                 470                 475                 480
Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln
                485                 490                 495
Val Pro Gln Asn Val Leu Ser Leu Leu Arg Leu Ser Leu Leu Thr Glu
            500                 505                 510
Val Glu Thr Pro Ile Arg Asn Glu Trp Gly Ser Arg Ser Asn Asp Ser
```

```
        515                 520                 525
Ser Asp Pro Leu Glu Ser Leu Leu Thr Glu Val Glu Thr Pro Ile Arg
    530                 535                 540
Asn Glu Trp Gly Ser Arg Ser Asn Asp Ser Ser Asp Pro Gly Ser Ser
545                 550                 555                 560
Leu Leu Thr Glu Val Glu Thr Pro Ile Arg Asn Glu Trp Gly Ser Arg
                565                 570                 575
Ser Asn Asp Ser Ser Asp Pro Glu Leu Ser Leu Leu Thr Glu Val Glu
            580                 585                 590
Thr Pro Ile Arg Asn Glu Trp Gly Ser Arg Ser Asn Asp Ser Ser Asp
        595                 600                 605
Pro Ser Arg
    610
```

<210> 529
<211> 1836
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2-4XM2e

<400> 529

```
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcagaa agcgtatgat 540
gtgaaagata cagcagtaac aacgaaagct tatgccaata atggtactac actggatgta 600
tcgggtcttg atgatgcagc tattaaagcg gctacgggtg gtacgaatgg tacggcttct 660
gtaaccggtg gtgcggttaa atttgacgca gataataaca agtactttgt tactattggt 720
ggctttactg gtgctgatgc cgccaaaaat ggcgattatg aagttaacgt tgctactgac 780
ggtacagtaa cccttgcggc tggcgcaact aaaaccacaa tgcctgctgg tgcgacaact 840
aaaacagaag tacaggagtt aaaagataca ccggcagttg tttcagcaga tgctaaaaat 900
gccttaattg ctggcggcgt tgacgctacc gatgctaatg gcgctgagtt ggtcaaaatg 960
tcttataccg ataaaaatgg taagacaatt gaaggcggtt atgcgcttaa agctggcgat 1020
aagtattacg ccgcagatta cgatgaagcg acaggagcaa ttaaagctaa aaccacaagt 1080
tatactgctg ctgacggcac taccaaaaca gcggctaacc aactgggtgg cgtagacggt 1140
aaaaccgaag tcgttactat cgacggtaaa acctacaatg ccagcaaagc cgctggtcat 1200
gatttcaaag cacaaccaga gctggcggaa gcagccgcta aaaccaccga aaacccgctg 1260
cagaaaattg atgccgcgct ggcgcaggtg gatgcgctgc gctctgatct gggtgcggta 1320
caaaaccgtt tcaactctgc tatcaccaac ctgggcaata ccgtaaacaa tctgtctgaa 1380
gcgcgtagcc gtatcgaaga ttccgactac gcgaccgaag tttccaacat gtctcgcgcg 1440
cagattttgc agcaggccgg tacttccgtt ctggcgcagg ctaaccaggt cccgcagaac 1500
gtgctgtctc tgttacgtct gagcttgctg actgaggttg agaccccgat tcgcaacgaa 1560
tggggttccc gttccaacga ttcttccgac ccgctcgaga gcctgttgac cgaggttgaa 1620
acccctatcc gtaatgaatg gggctcccgt agcaacgact cttctgaccc aggatcctcc 1680
ctcttgaccg aagtggaaac gcctattcgt aatgagtggg gttctcgtag caatgacagc 1740
tcggacccgg agctctcgct gctgacggaa gtggagactc cgatccgtaa cgagtggggc 1800
tctcgctcta acgatagctc agacccgtct agataa                          1836
```

<210> 530
<211> 19

<212> PRT
<213> Artificial Sequence

<220>
<223> HA

<400> 530
Pro Ala Lys Leu Leu Lys Glu Arg Gly Arg Arg Gly Ala Ile Ala Gly
1 5 10 15
Phe Leu Glu

<210> 531
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> HA

<400> 531
Glu Val Glu Thr Pro Ile Arg Asn Glu Trp
1 5 10

<210> 532
<211> 26
<212> PRT
<213> Artificial Sequence

<220>
<223> HA FluB

<400> 532
Ser Leu Trp Ser Glu Glu Pro Ala Lys Leu Leu Lys Glu Arg Gly Phe
1 5 10 15
Phe Gly Ala Ile Ala Gly Phe Leu Glu Glu
20 25

<210> 533
<211> 26
<212> PRT
<213> Artificial Sequence

<220>
<223> HA FluA H1/H0

<400> 533
Ser Leu Trp Ser Glu Glu Asn Ile Pro Ser Ile Gln Ser Arg Gly Leu
1 5 10 15
Phe Gly Ala Ile Ala Gly Phe Ile Glu Glu
20 25

<210> 534
<211> 26
<212> PRT
<213> Artificial Sequence

<220>
<223> HA FluA H3/H0

<400> 534
Ser Leu Trp Ser Glu Glu Asn Val Pro Glu Lys Gln Thr Arg Gly Ile
1 5 10 15
Phe Gly Ala Ile Ala Gly Phe Ile Glu Glu
20 25

<210> 535
<211> 30
<212> PRT
<213> Artificial Sequence

<220>
<223> HA Flu A H5/H0

<400> 535
Ser Leu Trp Ser Glu Glu Glu Trp Glu Glu Arg Glu Arg Arg Arg Lys
1 5 10 15
Lys Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Glu
20 25 30

<210> 536
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> M2

<400> 536
Ser Leu Leu Thr Glu Val Glu Thr Pro Thr Arg Asn Glu Trp Glu Ser
1 5 10 15
Arg Ser Ser Asp Ser Ser Asp Pro
20

<210> 537
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> M2

<400> 537

Ser Leu Leu Thr Glu Val Glu Thr Pro Thr Arg Asn Glu Trp Glu Cys
1 5 10 15
Arg Cys Ser Asp Ser Ser Asp Pro
20

<210> 538
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> M2

<400> 538
Ser Leu Leu Thr Glu Val Glu Thr Leu Thr Arg Asn Gly Trp Gly Ser
1 5 10 15
Arg Ser Ser Asp Ser Ser Asp Pro
20

<210> 539
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> M2

<400> 539
Ser Leu Leu Thr Glu Val Glu Thr Leu Thr Arg Asn Gly Trp Gly Cys
1 5 10 15
Arg Cys Ser Asp Ser Ser Asp Pro
20

<210> 540
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> M2

<400> 540
Ser Leu Leu Thr Glu Val Glu Thr Pro Thr Arg Asn Gly Trp Glu Ser
1 5 10 15
Lys Ser Ser Asp Ser Ser Asp Pro
20

<210> 541
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> M2

<400> 541
Ser Leu Leu Thr Glu Val Glu Thr Pro Thr Arg Asn Gly Trp Glu Cys
1 5 10 15
Lys Cys Ser Asp Ser Ser Asp Pro
20

<210> 542
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> M2

<400> 542
Ser Leu Leu Thr Glu Val Glu Thr Leu Thr Arg Asn Gly Trp Glu Ser
1 5 10 15
Lys Ser Arg Asp Ser Ser Asp Pro
20

<210> 543
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> M2

<400> 543
Ser Leu Leu Thr Glu Val Glu Thr Leu Thr Arg Asn Gly Trp Glu Cys
1 5 10 15
Lys Cys Arg Asp Ser Ser Asp Pro
20

<210> 544
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> M2

<400> 544
Ser Leu Leu Thr Glu Val Glu Thr Pro Ile Arg Asn Glu Trp Gly Ser
1 5 10 15
Arg Ser Asn Asp Ser Ser Asp Pro
20

<210> 545
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> M2

<400> 545
Ser Leu Leu Thr Glu Val Glu Thr Pro Ile Arg Asn Glu Trp Gly Cys
1 5 10 15
Arg Cys Asn Gly Ser Ser Asp Pro
20

<210> 546
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> M2

<400> 546
Ser Leu Leu Thr Glu Val Glu Thr Pro Ile Arg Asn Glu Trp Gly Cys
1 5 10 15
Lys Cys Asn Asp Ser Ser Asp Pro
20

<210> 547
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> M2

<400> 547
Ser Leu Leu Thr Glu Val Glu Thr Pro Ile Arg Asn Glu Trp Gly Cys
1 5 10 15
Arg Cys Asn Asp Ser Ser Asp Pro
20

<210> 548
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> M2

<400> 548
Ser Leu Leu Thr Glu Val Glu Thr Pro Ile Arg Asn Glu Trp Gly Cys

1 5 10 15
Arg Cys Asn Asp Ser Ser Asp Pro
20

<210> 549
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> M2

<400> 549
Ser Leu Leu Thr Glu Val Glu Thr Pro Ile Arg Asn Glu Trp Glu Cys
1 5 10 15
Arg Cys Asn Gly Ser Ser Asp Pro
20

<210> 550
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> M2

<400> 550
Ser Leu Leu Thr Glu Val Glu Thr Pro Thr Arg Asn Glu Trp Glu Cys
1 5 10 15
Arg Cys Ser Asp Ser Ser Asp Pro
20

<210> 551
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> M2

<400> 551
Ser Leu Leu Thr Glu Val Glu Thr Pro Ile Arg Asn Glu Trp Gly Ser
1 5 10 15
Arg Ser Asn Gly Ser Ser Asp Pro
20

<210> 552
<211> 24
<212> PRT
<213> Artificial Sequence

<220>

<223> M2

<400> 552
Ser Leu Leu Thr Glu Val Glu Thr Leu Thr Arg Asn Gly Trp Gly Cys
1 5 10 15
Arg Cys Ser Asp Ser Ser Asp Pro
20

<210> 553
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> M2

<400> 553
Ser Leu Leu Thr Glu Val Glu Thr Pro Thr Arg Asn Glu Trp Glu Cys
1 5 10 15
Arg Cys Ser Asp Ser Ser Asp Pro
20

<210> 554
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> M2

<400> 554
Ser Leu Leu Thr Glu Val Glu Thr Pro Thr Arg Asn Gly Trp Glu Cys
1 5 10 15
Lys Cys Ser Asp Ser Ser Asp Pro
20

<210> 555
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> M2

<400> 555
Ser Leu Leu Thr Glu Val Glu Thr Pro Thr Arg Asn Gly Trp Gly Cys
1 5 10 15
Arg Cys Ser Gly Ser Ser Asp Pro
20

<210> 556

<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> M2

<400> 556
Ser Leu Leu Thr Glu Val Glu Thr Leu Thr Arg Asn Gly Trp Glu Cys
1 5 10 15
Lys Cys Arg Asp Ser Ser Asp Pro
20

<210> 557
<211> 811
<212> PRT
<213> Artificial Sequence

<220>
<223> STF2delta.HA

<400> 557
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
1 5 10 15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
20 25 30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
35 40 45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
50 55 60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65 70 75 80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
85 90 95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
100 105 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
115 120 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
130 135 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145 150 155 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val His
165 170 175
Gly Ala Pro Val Asp Pro Ala Ser Pro Trp Thr Glu Asn Pro Leu Gln
180 185 190
Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala Leu Arg Ser Asp Leu
195 200 205
Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile Thr Asn Leu Gly Asn
210 215 220
Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg Ile Glu Asp Ser Asp
225 230 235 240
Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala Gln Ile Leu Gln Gln
245 250 255
Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln Val Pro Gln Asn Val
260 265 270

Leu Ser Leu Leu Arg Glu Phe Ser Arg Tyr Pro Ala Gln Trp Arg Pro
275 280 285
Leu Asp Gln Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Glu Gln
290 295 300
Val Asp Thr Ile Met Glu Lys Asn Val Thr Val Thr His Ala Gln Asp
305 310 315 320
Ile Leu Glu Lys Lys His Asn Gly Lys Leu Cys Asp Leu Asp Gly Val
325 330 335
Lys Pro Leu Ile Leu Arg Asp Cys Ser Val Ala Gly Trp Leu Leu Gly
340 345 350
Asn Pro Met Cys Asp Glu Phe Ile Asn Val Pro Glu Trp Ser Tyr Ile
355 360 365
Val Glu Lys Ala Asn Pro Val Asn Asp Leu Cys Tyr Pro Gly Asp Phe
370 375 380
Asn Asp Tyr Glu Glu Leu Lys His Leu Leu Ser Arg Ile Asn His Phe
385 390 395 400
Glu Lys Ile Gln Ile Ile Pro Lys Ser Ser Trp Ser Ser His Glu Ala
405 410 415
Ser Leu Gly Val Ser Ser Ala Cys Pro Tyr Gln Gly Lys Ser Ser Phe
420 425 430
Phe Arg Asn Val Val Trp Leu Ile Lys Lys Asn Ser Thr Tyr Pro Thr
435 440 445
Ile Lys Arg Ser Tyr Asn Asn Thr Asn Gln Glu Asp Leu Leu Val Leu
450 455 460
Trp Gly Ile His His Pro Asn Asp Ala Ala Glu Gln Thr Lys Leu Tyr
465 470 475 480
Gln Asn Pro Thr Thr Tyr Ile Ser Val Gly Thr Ser Thr Leu Asn Gln
485 490 495
Arg Leu Val Pro Arg Ile Ala Thr Arg Ser Lys Val Asn Gly Gln Ser
500 505 510
Gly Arg Met Glu Phe Phe Trp Thr Ile Leu Lys Pro Asn Asp Ala Ile
515 520 525
Asn Phe Glu Ser Asn Gly Asn Phe Ile Ala Pro Glu Tyr Ala Tyr Lys
530 535 540
Ile Val Lys Lys Gly Asp Ser Thr Ile Met Lys Ser Glu Leu Glu Tyr
545 550 555 560
Gly Asn Cys Asn Thr Lys Cys Gln Thr Pro Met Gly Ala Ile Asn Ser
565 570 575
Ser Met Pro Phe His Asn Ile His Pro Leu Thr Ile Gly Glu Cys Pro
580 585 590
Lys Tyr Val Lys Ser Asn Arg Leu Val Leu Ala Thr Gly Leu Arg Asn
595 600 605
Ser Pro Gln Arg Glu Arg Arg Arg Lys Lys Arg Gly Leu Phe Gly Ala
610 615 620
Ile Ala Gly Phe Ile Glu Gly Gly Trp Gln Gly Met Val Asp Gly Trp
625 630 635 640
Tyr Gly Tyr His His Ser Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp
645 650 655
Lys Glu Ser Thr Gln Lys Ala Ile Asp Gly Val Thr Asn Lys Val Asn
660 665 670
Ser Ile Ile Asp Lys Met Asn Thr Gln Phe Glu Ala Val Gly Arg Glu
675 680 685
Phe Asn Asn Leu Glu Arg Arg Ile Glu Asn Leu Asn Lys Lys Met Glu
690 695 700
Asp Gly Phe Leu Asp Val Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu
705 710 715 720
Met Glu Asn Glu Arg Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn
725 730 735
Leu Tyr Asp Lys Val Arg Leu Gln Leu Arg Asp Asn Ala Lys Glu Leu

```
            740                 745                 750
Gly Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp Asn Glu Cys Met
        755                 760                 765
Glu Ser Val Arg Asn Gly Thr Tyr Asp Tyr Pro Gln Tyr Ser Glu Glu
    770                 775                 780
Ala Arg Leu Lys Arg Glu Glu Ile Ser Gly Val Lys Leu Glu Ser Ile
785                 790                 795                 800
Gly Ile Tyr Gln Ile Leu Ser Ile Tyr Ser Thr
                805                 810
```

<210> 558
<211> 2439
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2delta.HA

<400> 558

```
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcatgg agcgccggtg 540
gatcctgcta gcccatggac cgaaaacccg ctgcagaaaa ttgatgccgc gctggcgcag 600
gtggatgcgc tgcgctctga tctgggtgcg gtacaaaacc gtttcaactc tgctatcacc 660
aacctgggca ataccgtaaa caatctgtct gaagcgcgta gccgtatcga agattccgac 720
tacgcgaccg aagtttccaa catgtctcgc gcgcagattt tgcagcaggc cggtacttcc 780
gttctggcgc aggctaacca ggtcccgcag aacgtgctgt ctctgttacg tgaattctct 840
agatatccag cacagtggcg gccgctcgac cagatctgta tcggttatca tgctaacaat 900
tctactgaac aagtagatac tatcatggag aagaacgtta cagttacaca tgcacaagat 960
atcctggaaa agaagcataa tggaaaactg tgtgaccttg atggtgttaa accactaata 1020
ttgcgtgact gctcagttgc tgggtggttg ttggggaatc caatgtgcga cgaatttatc 1080
aacgttccag aatggagtta cattgttgaa aaagctaacc ctgttaacga cttgtgttac 1140
ccaggcgatt ttaatgacta cgaggaactt aagcatttgt tgtcaagaat taaccacttc 1200
gagaaaattc aaattattcc aaagtcatct tggtcctccc atgaagcatc cctaggagtc 1260
tcttccgctt gcccttacca aggcaagagt tccttttttc gtaatgtcgt ctggctgatc 1320
aaaaagaact ccacctatcc aactataaag agatcataca acaacacaaa tcaggaggat 1380
ctgctagttc tgtggggcat tcaccacccc aatgacgcag ctgagcagac taaattgtac 1440
caaaacccaa ctacctatat atcagttggt acctcaactc ttaaccagcg actagtcccc 1500
cgtattgcta ctaggtcaaa ggttaatggt caaagtggac gaatggagtt tttctggact 1560
attttgaagc ccaacgatgc catcaacttc gaaagtaatg gaaatttcat agcccctgag 1620
tacgcttaca aaatcgttaa aaagggtgat tccactatca tgaaatctga actggaatac 1680
ggaaactgta acaccaaatg ccagacgcca atgggtgcca tcaactcttc tatgcctttt 1740
cacaacattc atcctttgac tattggtgaa tgcccaaagt acgtcaaatc taaccgtttg 1800
gtgttggcta ctggtctaag gaactccctt cagcgtgaaa gaagaagaaa gaagagggga 1860
ttattcggtg ctatcgctgg atttattgag ggaggatggc agggaatggt cgatggctgg 1920
tatggttacc atcactcaaa tgaacaggga agtggatacg cagctgataa agaatctact 1980
caaaaggcta tcgacggtgt tacaaacaag gtcaattcta ttatcgataa gatgaataca 2040
cagtttgagg ctgttggtag agagttcaat aatcttgaga gaagaatcga aaacctgaac 2100
aagaaaatgg aagacggatt tttagatgta tggacttaca atgctgagtt gttggtcttg 2160
atggagaatg aacgaacgtt ggacttccat gactccaatg tgaagaacct atatgacaaa 2220
gtgaggctgc aacttagaga caacgccaag gaattgggaa acgggtgctt cgagttttac 2280
```

cacaaatgcg acaacgaatg tatggaatca gtgagaaacg gtacctatga ttacccccaa 2340
tattccgagg aggcaagact gaagagagaa gagatatctg gtgtaaagtt ggaatccatc 2400
ggtatttatc agattctatc tatatattct acctaatag 2439

<210> 559
<211> 800
<212> PRT
<213> Artificial Sequence

<220>
<223> STF2delta.HA (PR8)

<400> 559

Met	Ala	Gln	Val	Ile	Asn	Thr	Asn	Ser	Leu	Ser	Leu	Leu	Thr	Gln	Asn
1				5				10					15		
Asn	Leu	Asn	Lys	Ser	Gln	Ser	Ala	Leu	Gly	Thr	Ala	Ile	Glu	Arg	Leu
			20					25					30		
Ser	Ser	Gly	Leu	Arg	Ile	Asn	Ser	Ala	Lys	Asp	Asp	Ala	Ala	Gly	Gln
		35					40					45			
Ala	Ile	Ala	Asn	Arg	Phe	Thr	Ala	Asn	Ile	Lys	Gly	Leu	Thr	Gln	Ala
	50					55					60				
Ser	Arg	Asn	Ala	Asn	Asp	Gly	Ile	Ser	Ile	Ala	Gln	Thr	Thr	Glu	Gly
65					70					75					80
Ala	Leu	Asn	Glu	Ile	Asn	Asn	Asn	Leu	Gln	Arg	Val	Arg	Glu	Leu	Ala
				85					90					95	
Val	Gln	Ser	Ala	Asn	Ser	Thr	Asn	Ser	Gln	Ser	Asp	Leu	Asp	Ser	Ile
			100					105					110		
Gln	Ala	Glu	Ile	Thr	Gln	Arg	Leu	Asn	Glu	Ile	Asp	Arg	Val	Ser	Gly
		115					120					125			
Gln	Thr	Gln	Phe	Asn	Gly	Val	Lys	Val	Leu	Ala	Gln	Asp	Asn	Thr	Leu
	130					135					140				
Thr	Ile	Gln	Val	Gly	Ala	Asn	Asp	Gly	Glu	Thr	Ile	Asp	Ile	Asp	Leu
145					150					155					160
Lys	Gln	Ile	Asn	Ser	Gln	Thr	Leu	Gly	Leu	Asp	Ser	Leu	Asn	Val	His
				165					170					175	
Gly	Ala	Pro	Val	Asp	Pro	Ala	Ser	Pro	Trp	Thr	Glu	Asn	Pro	Leu	Gln
			180					185					190		
Lys	Ile	Asp	Ala	Ala	Leu	Ala	Gln	Val	Asp	Ala	Leu	Arg	Ser	Asp	Leu
		195					200					205			
Gly	Ala	Val	Gln	Asn	Arg	Phe	Asn	Ser	Ala	Ile	Thr	Asn	Leu	Gly	Asn
	210					215					220				
Thr	Val	Asn	Asn	Leu	Ser	Glu	Ala	Arg	Ser	Arg	Ile	Glu	Asp	Ser	Asp
225					230					235					240
Tyr	Ala	Thr	Glu	Val	Ser	Asn	Met	Ser	Arg	Ala	Gln	Ile	Leu	Gln	Gln
				245					250					255	
Ala	Gly	Thr	Ser	Val	Leu	Ala	Gln	Ala	Asn	Gln	Val	Pro	Gln	Asn	Val
			260					265					270		
Leu	Ser	Leu	Leu	Arg	Glu	Phe	Ser	Arg	Tyr	Pro	Ala	Gln	Trp	Arg	Pro
		275					280					285			
Leu	Asp	Thr	Ile	Cys	Ile	Gly	Tyr	His	Ala	Asn	Asn	Ser	Thr	Asp	Thr
	290					295					300				
Val	Asp	Thr	Val	Leu	Glu	Lys	Asn	Val	Thr	Val	Thr	His	Ser	Val	Asn
305					310					315					320
Leu	Leu	Glu	Asp	Ser	His	Asn	Gly	Lys	Leu	Cys	Arg	Leu	Lys	Gly	Ile
				325					330					335	
Ala	Pro	Leu	Gln	Leu	Gly	Lys	Cys	Asn	Ile	Ala	Gly	Trp	Leu	Leu	Gly
			340					345					350		
Asn	Pro	Glu	Cys	Asp	Pro	Leu	Leu	Pro	Val	Arg	Ser	Trp	Ser	Tyr	Ile

```
        355                 360                 365
Val Glu Thr Pro Asn Ser Glu Asn Gly Ile Cys Tyr Pro Gly Asp Phe
    370                 375                 380
Ile Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe
385                 390                 395                 400
Glu Arg Phe Glu Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His Asn
                405                 410                 415
Thr Asn Gly Val Thr Ala Ala Cys Ser His Glu Gly Lys Ser Ser Phe
            420                 425                 430
Tyr Arg Asn Leu Leu Trp Leu Thr Glu Lys Glu Gly Ser Tyr Pro Lys
        435                 440                 445
Leu Lys Asn Ser Tyr Val Asn Lys Lys Gly Lys Glu Val Leu Val Leu
    450                 455                 460
Trp Gly Ile His His Pro Pro Asn Ser Lys Glu Gln Gln Asn Leu Tyr
465                 470                 475                 480
Gln Asn Glu Asn Ala Tyr Val Ser Val Val Thr Ser Asn Tyr Asn Arg
                485                 490                 495
Arg Phe Thr Pro Glu Ile Ala Glu Arg Pro Lys Val Arg Asp Gln Ala
            500                 505                 510
Gly Arg Met Asn Tyr Tyr Trp Thr Leu Leu Lys Pro Gly Asp Thr Ile
        515                 520                 525
Ile Phe Glu Ala Asn Gly Asn Leu Ile Ala Pro Met Tyr Ala Phe Ala
    530                 535                 540
Leu Ser Arg Gly Phe Gly Ser Gly Ile Ile Thr Ser Asn Ala Ser Met
545                 550                 555                 560
His Glu Cys Asn Thr Lys Cys Gln Thr Pro Leu Gly Ala Ile Asn Ser
                565                 570                 575
Ser Leu Pro Tyr Gln Asn Ile His Pro Val Thr Ile Gly Glu Cys Pro
            580                 585                 590
Lys Tyr Val Arg Ser Ala Lys Leu Arg Met Val Thr Gly Leu Arg Asn
        595                 600                 605
Ile Pro Ser Ile Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe
    610                 615                 620
Ile Glu Gly Gly Trp Thr Gly Met Ile Asp Gly Trp Tyr Gly Tyr His
625                 630                 635                 640
His Gln Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Gln Lys Ser Thr
                645                 650                 655
Gln Asn Ala Ile Asn Gly Ile Thr Asn Lys Val Asn Thr Val Ile Glu
            660                 665                 670
Lys Met Asn Ile Gln Phe Thr Ala Val Gly Lys Glu Phe Asn Lys Leu
        675                 680                 685
Glu Lys Arg Met Glu Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Leu
    690                 695                 700
Asp Ile Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu
705                 710                 715                 720
Arg Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys
                725                 730                 735
Val Lys Ser Gln Leu Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys
            740                 745                 750
Phe Glu Phe Tyr His Lys Cys Asp Asn Glu Cys Met Glu Ser Val Arg
        755                 760                 765
Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ser Lys Leu Asn
    770                 775                 780
Arg Glu Lys Val Asp Gly Val Lys Leu Glu Ser Met Gly Ile Tyr Gln
785                 790                 795                 800


<210> 560
<211> 2439
```

<212> DNA
<213> Artificial Sequence

<220>
<223> STF2delta.HA (PR8)

<400> 560
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcatgg agcgccggtg 540
gatcctgcta gcccatggac cgaaaacccg ctgcagaaaa ttgatgccgc gctggcgcag 600
gtggatgcgc tgcgctctga tctgggtgcg gtacaaaacc gtttcaactc tgctatcacc 660
aacctgggca ataccgtaaa caatctgtct gaagcgcgta gccgtatcga agattccgac 720
tacgcgaccg aagtttccaa catgtctcgc gcgcagattt tgcagcaggc cggtacttcc 780
gttctggcgc aggctaacca ggtcccgcag aacgtgctgt ctctgttacg tgaattctct 840
agatatccag cacagtggcg gccgctcgac cagatctgta tcggttatca tgctaacaat 900
tctactgaac aagtagatac tatcatggag aagaacgtta cagttacaca tgcacaagat 960
atcctggaaa agaagcataa tggaaaactg tgtgaccttg atggtgttaa accactaata 1020
ttgcgtgact gctcagttgc tgggtggttg ttggggaatc caatgtgcga cgaatttatc 1080
aacgttccag aatggagtta cattgttgaa aaagctaacc ctgttaacga cttgtgttac 1140
ccaggcgatt ttaatgacta cgaggaactt aagcatttgt tgtcaagaat taaccacttc 1200
gagaaaattc aaattattcc aaagtcatct tggtcctccc atgaagcatc cctaggagtc 1260
tcttccgctt gcccttacca aggcaagagt tccttttttc gtaatgtcgt ctggctgatc 1320
aaaaagaact ccacctatcc aactataaag agatcataca acaacacaaa tcaggaggat 1380
ctgctagttc tgtggggcat tcaccacccc aatgacgcag ctgagcagac taaattgtac 1440
caaaacccaa ctacctatat atcagttggt acctcaactc ttaaccagcg actagtcccc 1500
cgtattgcta ctaggtcaaa ggttaatggt caaagtggac gaatggagtt tttctggact 1560
attttgaagc ccaacgatgc catcaacttc gaaagtaatg gaaatttcat agcccctgag 1620
tacgcttaca aaatcgttaa aaagggtgat tccactatca tgaaatctga actggaatac 1680
ggaaactgta acaccaaatg ccagacgcca atgggtgcca tcaactcttc tatgcctttt 1740
cacaacattc atcctttgac tattggtgaa tgcccaaagt acgtcaaatc taaccgtttg 1800
gtgttggcta ctggtctaag gaactccctt cagcgtgaaa gaagaagaaa gaagagggga 1860
ttattcggtg ctatcgctgg atttattgag ggaggatggc agggaatggt cgatggctgg 1920
tatggttacc atcactcaaa tgaacaggga agtggatacg cagctgataa agaatctact 1980
caaaaggcta tcgacggtgt tacaaacaag gtcaattcta ttatcgataa gatgaataca 2040
cagtttgagg ctgttggtag agagttcaat aatcttgaga gaagaatcga aaacctgaac 2100
aagaaaatgg aagacggatt tttagatgta tggacttaca atgctgagtt gttggtcttg 2160
atggagaatg aacgaacgtt ggacttccat gactccaatg tgaagaacct atatgacaaa 2220
gtgaggctgc aacttagaga caacgccaag gaattgggaa acgggtgctt cgagttttac 2280
cacaaatgcg acaacgaatg tatggaatca gtgagaaacg gtacctatga ttacccccaa 2340
tattccgagg aggcaagact gaagagagaa gagatatctg gtgtaaagtt ggaatccatc 2400
ggtatttatc agattctatc tatatattct acctaatag 2439

<210> 561
<211> 511
<212> PRT
<213> Artificial Sequence

<220>
<223> HA (PR8)

<400> 561

```
Asp Thr Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Asp Thr Val
 1               5                  10                  15
Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His Ser Val Asn Leu
            20                  25                  30
Leu Glu Asp Ser His Asn Gly Lys Leu Cys Arg Leu Lys Gly Ile Ala
        35                  40                  45
Pro Leu Gln Leu Gly Lys Cys Asn Ile Ala Gly Trp Leu Leu Gly Asn
    50                  55                  60
Pro Glu Cys Asp Pro Leu Leu Pro Val Arg Ser Trp Ser Tyr Ile Val
65                  70                  75                  80
Glu Thr Pro Asn Ser Glu Asn Gly Ile Cys Tyr Pro Gly Asp Phe Ile
                85                  90                  95
Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe Glu
            100                 105                 110
Arg Phe Glu Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His Asn Thr
        115                 120                 125
Asn Gly Val Thr Ala Ala Cys Ser His Glu Gly Lys Ser Ser Phe Tyr
    130                 135                 140
Arg Asn Leu Leu Trp Leu Thr Glu Lys Glu Gly Ser Tyr Pro Lys Leu
145                 150                 155                 160
Lys Asn Ser Tyr Val Asn Lys Lys Gly Lys Glu Val Leu Val Leu Trp
                165                 170                 175
Gly Ile His His Pro Pro Asn Ser Lys Glu Gln Gln Asn Leu Tyr Gln
            180                 185                 190
Asn Glu Asn Ala Tyr Val Ser Val Val Thr Ser Asn Tyr Asn Arg Arg
        195                 200                 205
Phe Thr Pro Glu Ile Ala Glu Arg Pro Lys Val Arg Asp Gln Ala Gly
    210                 215                 220
Arg Met Asn Tyr Tyr Trp Thr Leu Leu Lys Pro Gly Asp Thr Ile Ile
225                 230                 235                 240
Phe Glu Ala Asn Gly Asn Leu Ile Ala Pro Met Tyr Ala Phe Ala Leu
                245                 250                 255
Ser Arg Gly Phe Gly Ser Gly Ile Ile Thr Ser Asn Ala Ser Met His
            260                 265                 270
Glu Cys Asn Thr Lys Cys Gln Thr Pro Leu Gly Ala Ile Asn Ser Ser
        275                 280                 285
Leu Pro Tyr Gln Asn Ile His Pro Val Thr Ile Gly Glu Cys Pro Lys
    290                 295                 300
Tyr Val Arg Ser Ala Lys Leu Arg Met Val Thr Gly Leu Arg Asn Ile
305                 310                 315                 320
Pro Ser Ile Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile
                325                 330                 335
Glu Gly Gly Trp Thr Gly Met Ile Asp Gly Trp Tyr Gly Tyr His His
            340                 345                 350
Gln Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Gln Lys Ser Thr Gln
        355                 360                 365
Asn Ala Ile Asn Gly Ile Thr Asn Lys Val Asn Thr Val Ile Glu Lys
    370                 375                 380
Met Asn Ile Gln Phe Thr Ala Val Gly Lys Glu Phe Asn Lys Leu Glu
385                 390                 395                 400
Lys Arg Met Glu Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Leu Asp
                405                 410                 415
Ile Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu Arg
            420                 425                 430
Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys Val
        435                 440                 445
Lys Ser Gln Leu Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys Phe
    450                 455                 460
```

```
Glu Phe Tyr His Lys Cys Asp Asn Glu Cys Met Glu Ser Val Arg Asn
465                 470                 475                 480
Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ser Lys Leu Asn Arg
                485                 490                 495
Glu Lys Val Asp Gly Val Lys Leu Glu Ser Met Gly Ile Tyr Gln
            500                 505                 510
```

<210> 562
<211> 1572
<212> DNA
<213> Artificial Sequence

<220>
<223> HA (PR8)

<400> 562
gaccagatct gtatcggtta tcatgctaac aattctactg aacaagtaga tactatcatg 60
gagaagaacg ttacagttac acatgcacaa gatatcctgg aaaagaagca taatggaaaa 120
ctgtgtgacc ttgatggtgt taaaccacta atattgcgtg actgctcagt tgctgggtgg 180
ttgttgggga atccaatgtg cgacgaattt atcaacgttc cagaatggag ttacattgtt 240
gaaaaagcta accctgttaa cgacttgtgt tacccaggcg attttaatga ctacgaggaa 300
cttaagcatt tgttgtcaag aattaaccac ttcgagaaaa ttcaaattat tccaaagtca 360
tcttggtcct cccatgaagc atcctagga gtctcttccg cttgccctta ccaaggcaag 420
agttcctttt ttcgtaatgt cgtctggctg atcaaaaaga actccaccta tccaactata 480
aagagatcat acaacaacac aaatcaggag gatctgctag ttctgtgggg cattcaccac 540
cccaatgacg cagctgagca gactaaattg taccaaaacc caactaccta tatatcagtt 600
ggtacctcaa ctcttaacca gcgactagtc ccccgtattg ctactaggtc aaaggttaat 660
ggtcaaagtg gacgaatgga gtttttctgg actattttga agcccaacga tgccatcaac 720
ttcgaaagta atggaaattt catagcccct gagtacgctt acaaaatcgt taaaaagggt 780
gattccacta tcatgaaatc tgaactggaa tacggaaact gtaacaccaa atgccagacg 840
ccaatgggtg ccatcaactc ttctatgcct tttcacaaca ttcatccttt gactattggt 900
gaatgcccaa agtacgtcaa atctaaccgt ttggtgttgg ctactggtct aaggaactcc 960
cctcagcgtg aaagaagaag aaagaagagg ggattattcg gtgctatcgc tggatttatt 1020
gagggaggat ggcagggaat ggtcgatggc tggtatggtt accatcactc aaatgaacag 1080
ggaagtggat acgcagctga taaagaatct actcaaaagg ctatcgacgg tgttacaaac 1140
aaggtcaatt ctattatcga taagatgaat acacagtttg aggctgttgg tagagagttc 1200
aataatcttg agagaagaat cgaaaacctg aacaagaaaa tggaagacgg atttttagat 1260
gtatggactt acaatgctga gttgttggtc ttgatggaga atgaacgaac gttggacttc 1320
catgactcca atgtgaagaa cctatatgac aaagtgaggc tgcaacttag agacaacgcc 1380
aaggaattgg gaaacgggtg cttcgagttt taccacaaat gcgacaacga atgtatggaa 1440
tcagtgagaa acggtaccta tgattacccc caatattccg aggaggcaag actgaagaga 1500
gaagagatat ctggtgtaaa gttggaatcc atcggtattt atcagattct atctatatat 1560
tctacctaat ag 1572

<210> 563
<211> 261
<212> PRT
<213> E. coli

<400> 563
```
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Ile Thr Gln Asn
 1               5                  10                  15
Asn Ile Asn Lys Asn Gln Ser Ala Leu Ser Ser Ser Ile Glu Arg Leu
            20                  25                  30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
        35                  40                  45
Ala Ile Ala Asn Arg Phe Thr Ser Asn Ile Lys Gly Leu Thr Gln Ala
```

```
    50                  55                  60
Ala Arg Asn Ala Asn Asp Gly Ile Ser Val Ala Gln Thr Thr Glu Gly
65                  70                  75                  80
Ala Leu Ser Glu Ile Asn Asn Asn Leu Gln Arg Ile Arg Glu Leu Thr
                85                  90                  95
Val Gln Ala Ser Thr Gly Thr Asn Ser Asp Ser Asp Leu Asp Ser Ile
            100                 105                 110
Gln Asp Glu Ile Lys Ser Arg Leu Asp Glu Ile Asp Arg Val Ser Gly
        115                 120                 125
Gln Thr Gln Phe Asn Gly Val Asn Val Leu Ala Lys Asp Gly Ser Met
    130                 135                 140
Lys Ile Gln Val Gly Ala Asn Asp Gly Gln Thr Ile Thr Ile Asp Leu
145                 150                 155                 160
Lys Lys Ile Asp Ser Asp Thr Leu Gly Thr Lys Asp Pro Leu Lys Ala
                165                 170                 175
Leu Asp Glu Ala Ile Ser Ser Ile Asp Lys Phe Arg Ser Ser Leu Gly
            180                 185                 190
Ala Ile Gln Asn Arg Leu Asp Ser Ala Val Thr Asn Leu Asn Asn Thr
        195                 200                 205
Thr Thr Asn Leu Ser Glu Ala Gln Ser Arg Ile Gln Asp Ala Asp Tyr
    210                 215                 220
Ala Thr Glu Val Ser Asn Met Ser Lys Ala Gln Ile Ile Gln Gln Ala
225                 230                 235                 240
Gly Asn Ser Val Leu Ala Lys Ala Asn Gln Val Pro Gln Gln Val Leu
                245                 250                 255
Ser Leu Leu Gln Gly
            260
```

<210> 564
<211> 522
<212> PRT
<213> Artificial Sequence

<220>
<223> H5N1 HA

<400> 564

```
Asp Gln Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Glu Gln Val
 1                  5                  10                  15
Asp Thr Ile Met Glu Lys Asn Val Thr Val Thr His Ala Gln Asp Ile
            20                  25                  30
Leu Glu Lys Lys His Asn Gly Lys Leu Cys Asp Leu Asp Gly Val Lys
        35                  40                  45
Pro Leu Ile Leu Arg Asp Cys Ser Val Ala Gly Trp Leu Leu Gly Asn
    50                  55                  60
Pro Met Cys Asp Glu Phe Ile Asn Val Pro Glu Trp Ser Tyr Ile Val
65                  70                  75                  80
Glu Lys Ala Asn Pro Val Asn Asp Leu Cys Tyr Pro Gly Asp Phe Asn
                85                  90                  95
Asp Tyr Glu Glu Leu Lys His Leu Leu Ser Arg Ile Asn His Phe Glu
            100                 105                 110
Lys Ile Gln Ile Ile Pro Lys Ser Ser Trp Ser Ser His Glu Ala Ser
        115                 120                 125
Leu Gly Val Ser Ser Ala Cys Pro Tyr Gln Gly Lys Ser Ser Phe Phe
    130                 135                 140
Arg Asn Val Val Trp Leu Ile Lys Lys Asn Ser Thr Tyr Pro Thr Ile
145                 150                 155                 160
```

```
Lys Arg Ser Tyr Asn Asn Thr Asn Gln Glu Asp Leu Leu Val Leu Trp
                165                 170                 175
Gly Ile His His Pro Asn Asp Ala Ala Glu Gln Thr Lys Leu Tyr Gln
            180                 185                 190
Asn Pro Thr Thr Tyr Ile Ser Val Gly Thr Ser Thr Leu Asn Gln Arg
        195                 200                 205
Leu Val Pro Arg Ile Ala Thr Arg Ser Lys Val Asn Gly Gln Ser Gly
    210                 215                 220
Arg Met Glu Phe Phe Trp Thr Ile Leu Lys Pro Asn Asp Ala Ile Asn
225                 230                 235                 240
Phe Glu Ser Asn Gly Asn Phe Ile Ala Pro Glu Tyr Ala Tyr Lys Ile
                245                 250                 255
Val Lys Lys Gly Asp Ser Thr Ile Met Lys Ser Glu Leu Glu Tyr Gly
            260                 265                 270
Asn Cys Asn Thr Lys Cys Gln Thr Pro Met Gly Ala Ile Asn Ser Ser
        275                 280                 285
Met Pro Phe His Asn Ile His Pro Leu Thr Ile Gly Glu Cys Pro Lys
    290                 295                 300
Tyr Val Lys Ser Asn Arg Leu Val Leu Ala Thr Gly Leu Arg Asn Ser
305                 310                 315                 320
Pro Gln Arg Glu Arg Arg Arg Lys Lys Arg Gly Leu Phe Gly Ala Ile
                325                 330                 335
Ala Gly Phe Ile Glu Gly Gly Trp Gln Gly Met Val Asp Gly Trp Tyr
            340                 345                 350
Gly Tyr His His Ser Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Lys
        355                 360                 365
Glu Ser Thr Gln Lys Ala Ile Asp Gly Val Thr Asn Lys Val Asn Ser
    370                 375                 380
Ile Ile Asp Lys Met Asn Thr Gln Phe Glu Ala Val Gly Arg Glu Phe
385                 390                 395                 400
Asn Asn Leu Glu Arg Arg Ile Glu Asn Leu Asn Lys Lys Met Glu Asp
                405                 410                 415
Gly Phe Leu Asp Val Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Met
            420                 425                 430
Glu Asn Glu Arg Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu
        435                 440                 445
Tyr Asp Lys Val Arg Leu Gln Leu Arg Asp Asn Ala Lys Glu Leu Gly
    450                 455                 460
Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp Asn Glu Cys Met Glu
465                 470                 475                 480
Ser Val Arg Asn Gly Thr Tyr Asp Tyr Pro Gln Tyr Ser Glu Glu Ala
                485                 490                 495
Arg Leu Lys Arg Glu Glu Ile Ser Gly Val Lys Leu Glu Ser Ile Gly
            500                 505                 510
Ile Tyr Gln Ile Leu Ser Ile Tyr Ser Thr
        515                 520
```

<210> 565
<211> 1572
<212> DNA
<213> Artificial Sequence

<220>
<223> H5N1 HA

<400> 565
gaccagatct gtatcggtta tcatgctaac aattctactg aacaagtaga tactatcatg 60

gagaagaacg ttacagttac acatgcacaa gatatcctgg aaaagaagca taatggaaaa 120
ctgtgtgacc ttgatggtgt taaaccacta atattgcgtg actgctcagt tgctgggtgg 180
ttgttgggga atccaatgtg cgacgaattt atcaacgttc cagaatggag ttacattgtt 240
gaaaaagcta accctgttaa cgacttgtgt tacccaggcg attttaatga ctacgaggaa 300
cttaagcatt tgttgtcaag aattaaccac ttcgagaaaa ttcaaattat tccaaagtca 360
tcttggtcct cccatgaagc atccctagga gtctcttccg cttgccctta ccaaggcaag 420
agttcctttt ttcgtaatgt cgtctggctg atcaaaaaga actccaccta tccaactata 480
aagagatcat acaacaacac aaatcaggag gatctgctag ttctgtgggg cattcaccac 540
cccaatgacg cagctgagca gactaaattg taccaaaacc caactaccta tatatcagtt 600
ggtacctcaa ctcttaacca gcgactagtc ccccgtattg ctactaggtc aaaggttaat 660
ggtcaaagtg gacgaatgga gtttttctgg actattttga agcccaacga tgccatcaac 720
ttcgaaagta atggaaattt catagcccct gagtacgctt acaaaatcgt taaaaagggt 780
gattccacta tcatgaaatc tgaactggaa tacggaaact gtaacaccaa atgccagacg 840
ccaatgggtg ccatcaactc ttctatgcct tttcacaaca ttcatccttt gactattggt 900
gaatgcccaa agtacgtcaa atctaaccgt ttggtgttgg ctactggtct aaggaactcc 960
cctcagcgtg aaagaagaag aaagaagagg ggattattcg gtgctatcgc tggatttatt 1020
gagggaggat ggcagggaat ggtcgatggc tggtatggtt accatcactc aaatgaacag 1080
ggaagtggat acgcagctga taaagaatct actcaaaagg ctatcgacgg tgttacaaac 1140
aaggtcaatt ctattatcga taagatgaat acacagtttg aggctgttgg tagagagttc 1200
aataatcttg agagaagaat cgaaaacctg aacaagaaaa tggaagacgg atttttagat 1260
gtatggactt acaatgctga gttgttggtc ttgatggaga atgaacgaac gttggacttc 1320
catgactcca atgtgaagaa cctatatgac aaagtgaggc tgcaacttag agacaacgcc 1380
aaggaattgg gaaacgggtg cttcgagttt taccacaaat gcgacaacga atgtatggaa 1440
tcagtgagaa acggtaccta tgattacccc caatattccg aggaggcaag actgaagaga 1500
gaagagatat ctggtgtaaa gttggaatcc atcggtattt atcagattct atctatatat 1560
tctacctaat ag 1572

<210> 566
<211> 43
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<220>
<223> m = a or c

<400> 566
cgctcttcam tgagcttgct gactgaggtt gagaccccga ttc 43

<210> 567
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 567
cgctcttcac gcttattatc tagacgggtc tgagctatcg ttagagcgag 50

<210> 568
<211> 41
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 568
gctcttcagc ggctgagcaa taactagcat aacccttgg g 41

<210> 569
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 569
cgctcttcac agacgtaaca gagacagcac gttctgcgg 39

<210> 570
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> M2

<400> 570
Ser Leu Leu Thr Glu Val Glu Thr Pro Ile Arg Asn Glu Trp Gly Ser
1 5 10 15
Lys Ser Asn Asp Ser Ser Asp Pro
20

<210> 571
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> M2

<400> 571
Ser Leu Leu Thr Glu Val Glu Thr Pro Ile Arg Asn Glu Trp Gly Ser
1 5 10 15
Arg Ser Asn Asp Ser Ser Asp Pro
20

<210> 572
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> M2

<400> 572
Ser Leu Leu Thr Glu Val Glu Thr Pro Ile Arg Asn Glu Trp Glu Ser
1 5 10 15
Arg Ser Asn Gly Ser Ser Asp Pro
20

<210> 573
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> M2

<400> 573
Ser Leu Leu Thr Glu Val Glu Thr Pro Thr Arg Asn Glu Trp Glu Ser
1 5 10 15
Arg Ser Ser Asp Ser Ser Asp Pro
20

<210> 574
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> M2

<400> 574
Ser Leu Leu Thr Glu Val Glu Thr Leu Thr Arg Asn Gly Trp Gly Ser
1 5 10 15
Arg Ser Ser Asp Ser Ser Asp Pro
20

<210> 575
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> M2

<400> 575
Ser Leu Leu Thr Glu Val Glu Thr Pro Thr Arg Asn Glu Trp Glu Ser
1 5 10 15
Arg Ser Ser Asp Ser Ser Asp Pro
20

<210> 576
<211> 24
<212> PRT

<213> Artificial Sequence

<220>
<223> M2

<400> 576
Ser Leu Leu Thr Glu Val Glu Thr Pro Thr Arg Asn Gly Trp Glu Ser
1 5 10 15
Lys Ser Ser Asp Ser Ser Asp Pro
20

<210> 577
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> M2

<400> 577
Ser Leu Leu Thr Glu Val Glu Thr Pro Thr Arg Asn Gly Trp Gly Ser
1 5 10 15
Arg Ser Ser Gly Ser Ser Asp Pro
20

<210> 578
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> M2

<400> 578
Ser Leu Leu Thr Glu Val Glu Thr Leu Thr Arg Asn Gly Trp Glu Ser
1 5 10 15
Lys Ser Arg Asp Ser Ser Asp Pro
20

<210> 579
<211> 641
<212> PRT
<213> Artificial Sequence

<220>
<223> pMT/STF2.4XM2e (H1)

<400> 579
Met Lys Leu Cys Ile Leu Leu Ala Val Val Ala Phe Val Gly Leu Ser
1 5 10 15
Leu Gly Arg Ser Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu
20 25 30

```
Thr Gln Asn Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile
        35                  40                  45
Glu Arg Leu Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala
    50                  55                  60
Ala Gly Gln Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu
65                  70                  75                  80
Thr Gln Ala Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr
                85                  90                  95
Thr Glu Gly Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg
            100                 105                 110
Glu Leu Ala Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu
        115                 120                 125
Asp Ser Ile Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg
    130                 135                 140
Val Ser Gly Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp
145                 150                 155                 160
Asn Thr Leu Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp
                165                 170                 175
Ile Asp Leu Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu
            180                 185                 190
Asn Val Gln Lys Ala Tyr Asp Val Lys Asp Thr Ala Val Thr Thr Lys
        195                 200                 205
Ala Tyr Ala Asn Asn Gly Thr Thr Leu Asp Val Ser Gly Leu Asp Asp
    210                 215                 220
Ala Ala Ile Lys Ala Ala Thr Gly Gly Thr Asn Gly Thr Ala Ser Val
225                 230                 235                 240
Thr Gly Gly Ala Val Lys Phe Asp Ala Asp Asn Asn Lys Tyr Phe Val
                245                 250                 255
Thr Ile Gly Gly Phe Thr Gly Ala Asp Ala Ala Lys Asn Gly Asp Tyr
            260                 265                 270
Glu Val Asn Val Ala Thr Asp Gly Thr Val Thr Leu Ala Ala Gly Ala
        275                 280                 285
Thr Lys Thr Thr Met Pro Ala Gly Ala Thr Thr Lys Thr Glu Val Gln
    290                 295                 300
Glu Leu Lys Asp Thr Pro Ala Val Val Ser Ala Asp Ala Lys Asn Ala
305                 310                 315                 320
Leu Ile Ala Gly Gly Val Asp Ala Thr Asp Ala Asn Gly Ala Glu Leu
                325                 330                 335
Val Lys Met Ser Tyr Thr Asp Lys Asn Gly Lys Thr Ile Glu Gly Gly
            340                 345                 350
Tyr Ala Leu Lys Ala Gly Asp Lys Tyr Tyr Ala Ala Asp Tyr Asp Glu
        355                 360                 365
Ala Thr Gly Ala Ile Lys Ala Lys Thr Thr Ser Tyr Thr Ala Ala Asp
    370                 375                 380
Gly Thr Thr Lys Thr Ala Ala Asn Gln Leu Gly Gly Val Asp Gly Lys
385                 390                 395                 400
Thr Glu Val Val Thr Ile Asp Gly Lys Thr Tyr Asn Ala Ser Lys Ala
                405                 410                 415
Ala Gly His Asp Phe Lys Ala Gln Pro Glu Leu Ala Glu Ala Ala Ala
            420                 425                 430
Lys Thr Thr Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala Gln
        435                 440                 445
Val Asp Ala Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe Asn
    450                 455                 460
Ser Ala Ile Thr Asn Leu Gly Asn Thr Val Asn Asn Leu Ser Glu Ala
465                 470                 475                 480
Arg Ser Arg Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn Met
                485                 490                 495
Ser Arg Ala Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala Gln
```

```
            500                 505                 510
Ala Asn Gln Val Pro Gln Asn Val Leu Ser Leu Leu Arg Lys Gly Asn
        515                 520                 525
Ser Lys Leu Glu Gly Gln Leu Glu Phe Ser Leu Leu Thr Glu Val Glu
    530                 535                 540
Thr Pro Ile Arg Asn Glu Trp Gly Ser Arg Ser Asn Asp Ser Ser Asp
545                 550                 555                 560
Pro Leu Glu Ser Leu Leu Thr Glu Val Glu Thr Pro Ile Arg Asn Glu
                565                 570                 575
Trp Gly Ser Arg Ser Asn Asp Ser Ser Asp Pro Gly Ser Ser Leu Leu
            580                 585                 590
Thr Glu Val Glu Thr Pro Ile Arg Asn Glu Trp Gly Ser Arg Ser Asn
        595                 600                 605
Asp Ser Ser Asp Pro Glu Leu Ser Leu Leu Thr Glu Val Glu Thr Pro
    610                 615                 620
Ile Arg Asn Glu Trp Gly Ser Arg Ser Asn Asp Ser Ser Asp Pro Ser
625                 630                 635                 640
Arg
```

<210> 580
<211> 1926
<212> DNA
<213> Artificial Sequence

<220>
<223> pMT/STF2.4XM2e (H1)

<400> 580

```
atgaagttat gcatattact ggccgtcgtg gcctttgttg gcctctcgct cgggagatct 60
gcacaagtaa tcaacactaa cagtctgtcg ctgctgaccc agaataacct gaacaaatcc 120
cagtccgcac tgggcaccgc tatcgagcgt ctgtcttctg gtctgcgtat caacagcgcg 180
aaagacgatg cggcaggtca ggcgattgct aaccgtttca ccgcgaacat caaaggtctg 240
actcaggctt cccgtaacgc taacgacggt atctccattg cgcagaccac tgaaggcgcg 300
ctgaacgaaa tcaacaacaa cctgcagcgt gtgcgtgaac tggcggttca gtctgctaac 360
agcaccaact cccagtctga cctcgactcc atccaggctg aaatcaccca gcgcctgaac 420
gaaatcgacc gtgtatccgg ccagactcag ttcaacggcg tgaaagtcct ggcgcaggac 480
aacaccctga ccatccaggt tggcgccaac gacggtgaaa ctatcgatat cgatctgaag 540
cagatcaact ctcagaccct gggtctggac tcactgaacg tgcagaaagc gtatgatgtg 600
aaagatacag cagtaacaac gaaagcttat gccaataatg gtactacact ggatgtatcg 660
ggtcttgatg atgcagctat taaagcggct acgggtggta cgaatggtac ggcttctgta 720
accggtggtg cggttaaatt tgacgcagat aataacaagt actttgttac tattggtggc 780
tttactggtg ctgatgccgc caaaaatggc gattatgaag ttaacgttgc tactgacggt 840
acagtaaccc ttgcggctgg cgcaactaaa accacaatgc ctgctggtgc gacaactaaa 900
acagaagtac aggagttaaa agatacaccg gcagttgttt cagcagatgc taaaaatgcc 960
ttaattgctg gcggcgttga cgctaccgat gctaatggcg ctgagttggt caaaatgtct 1020
tataccgata aaaatggtaa gacaattgaa ggcggttatg cgcttaaagc tggcgataag 1080
tattacgccg cagattacga tgaagcgaca ggagcaatta aagctaaaac tacaagttat 1140
actgctgctg acggcactac caaaacagcg gctaaccaac tgggtggcgt agacggtaaa 1200
accgaagtcg ttactatcga cggtaaaacc tacaatgcca gcaaagccgc tggtcatgat 1260
ttcaaagcac aaccagagct ggcggaagca gccgctaaaa ccaccgaaaa cccgctgcag 1320
aaaattgatg ccgcgctggc gcaggtggat gcgctgcgct ctgatctggg tgcggtacaa 1380
aaccgtttca actctgctat caccaacctg ggcaataccg taaacaatct gtctgaagcg 1440
cgtagccgta tcgaagattc cgactacgcg accgaagttt ccaacatgtc tcgcgcgcag 1500
attctgcagc aggccggtac ttccgttctg gcgcaggcta accaggtccc gcagaacgtg 1560
ctgtctctgt tacgtaaggg caattcgaag cttgaaggtc aattggaatt cagcttgctg 1620
actgaggttg agaccccgat tcgcaacgaa tggggttccc gttccaacga ttcttccgac 1680
```

ccgctcgaga gcctgttgac cgaggttgaa acccctatcc gtaatgaatg gggctcccgt 1740
agcaacgact cttctgaccc aggatcctcc ctcttgaccg aagtggaaac gcctattcgt 1800
aatgagtggg gttctcgtag caatgacagc tcggacccgg agctctcgct gctgacggaa 1860
gtggagactc cgatccgtaa cgagtggggc tctcgctcta acgatagctc agacccgtct 1920
agataa 1926

<210> 581
<211> 641
<212> PRT
<213> Artificial Sequence

<220>
<223> pMT/STF2.4XM2e (H5)

<400> 581

```
Met Lys Leu Cys Ile Leu Leu Ala Val Val Ala Phe Val Gly Leu Ser
 1               5                  10                  15
Leu Gly Arg Ser Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu
            20                  25                  30
Thr Gln Asn Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile
        35                  40                  45
Glu Arg Leu Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala
    50                  55                  60
Ala Gly Gln Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu
65                  70                  75                  80
Thr Gln Ala Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr
                85                  90                  95
Thr Glu Gly Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg
            100                 105                 110
Glu Leu Ala Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu
        115                 120                 125
Asp Ser Ile Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg
    130                 135                 140
Val Ser Gly Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp
145                 150                 155                 160
Asn Thr Leu Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp
                165                 170                 175
Ile Asp Leu Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu
            180                 185                 190
Asn Val Gln Lys Ala Tyr Asp Val Lys Asp Thr Ala Val Thr Thr Lys
        195                 200                 205
Ala Tyr Ala Asn Asn Gly Thr Thr Leu Asp Val Ser Gly Leu Asp Asp
    210                 215                 220
Ala Ala Ile Lys Ala Ala Thr Gly Gly Thr Asn Gly Thr Ala Ser Val
225                 230                 235                 240
Thr Gly Gly Ala Val Lys Phe Asp Ala Asp Asn Asn Lys Tyr Phe Val
                245                 250                 255
Thr Ile Gly Gly Phe Thr Gly Ala Asp Ala Ala Lys Asn Gly Asp Tyr
            260                 265                 270
Glu Val Asn Val Ala Thr Asp Gly Thr Val Thr Leu Ala Ala Gly Ala
        275                 280                 285
Thr Lys Thr Thr Met Pro Ala Gly Ala Thr Thr Lys Thr Glu Val Gln
    290                 295                 300
Glu Leu Lys Asp Thr Pro Ala Val Val Ser Ala Asp Ala Lys Asn Ala
305                 310                 315                 320
Leu Ile Ala Gly Gly Val Asp Ala Thr Asp Ala Asn Gly Ala Glu Leu
                325                 330                 335
Val Lys Met Ser Tyr Thr Asp Lys Asn Gly Lys Thr Ile Glu Gly Gly
```

```
            340                 345                 350
Tyr Ala Leu Lys Ala Gly Asp Lys Tyr Tyr Ala Ala Asp Tyr Asp Glu
        355                 360                 365
Ala Thr Gly Ala Ile Lys Ala Lys Thr Thr Ser Tyr Thr Ala Ala Asp
    370                 375                 380
Gly Thr Thr Lys Thr Ala Ala Asn Gln Leu Gly Gly Val Asp Gly Lys
385                 390                 395                 400
Thr Glu Val Val Thr Ile Asp Gly Lys Thr Tyr Asn Ala Ser Lys Ala
                405                 410                 415
Ala Gly His Asp Phe Lys Ala Gln Pro Glu Leu Ala Glu Ala Ala Ala
            420                 425                 430
Lys Thr Thr Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala Gln
        435                 440                 445
Val Asp Ala Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe Asn
    450                 455                 460
Ser Ala Ile Thr Asn Leu Gly Asn Thr Val Asn Asn Leu Ser Glu Ala
465                 470                 475                 480
Arg Ser Arg Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn Met
                485                 490                 495
Ser Arg Ala Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala Gln
            500                 505                 510
Ala Asn Gln Val Pro Gln Asn Val Leu Ser Leu Leu Arg Lys Gly Asn
        515                 520                 525
Ser Lys Leu Glu Gly Gln Leu Glu Phe Ser Leu Leu Thr Glu Val Glu
    530                 535                 540
Thr Pro Thr Arg Asn Glu Trp Glu Ser Arg Ser Ser Asp Ser Ser Asp
545                 550                 555                 560
Pro Leu Glu Ser Leu Leu Thr Glu Val Glu Thr Pro Thr Arg Asn Glu
                565                 570                 575
Trp Glu Ser Arg Ser Ser Asp Ser Ser Asp Pro Glu Ser Ser Leu Leu
            580                 585                 590
Thr Glu Val Glu Thr Pro Thr Arg Asn Glu Trp Glu Ser Arg Ser Ser
        595                 600                 605
Asp Ser Ser Asp Pro Gly Ser Ser Leu Leu Thr Glu Val Glu Thr Pro
    610                 615                 620
Thr Arg Asn Glu Trp Glu Ser Arg Ser Ser Asp Ser Ser Asp Pro Ser
625                 630                 635                 640
Arg
```

<210> 582
<211> 1927
<212> DNA
<213> Artificial Sequence

<220>
<223> pMT/STF2.4XM2e (H5)

<400> 582
atgaagttat gcatattact ggccgtcgtg gcctttgttg gcctctcgct cgggagatct 60
gcacaagtaa tcaacactaa cagtctgtcg ctgctgaccc agaataacct gaacaaatcc 120
cagtccgcac tgggcaccgc tatcgagcgt ctgtcttctg gtctgcgtat caacagcgcg 180
aaagacgatg cggcaggtca ggcgattgct aaccgtttca ccgcgaacat caaaggtctg 240
actcaggctt cccgtaacgc taacgacggt atctccattg cgcagaccac tgaaggcgcg 300
ctgaacgaaa tcaacaacaa cctgcagcgt gtgcgtgaac tggcggttca gtctgctaac 360
agcaccaact cccagtctga cctcgactcc atccaggctg aaatcaccca gcgcctgaac 420
gaaatcgacc gtgtatccgg ccagactcag ttcaacggcg tgaaagtcct ggcgcaggac 480

aacaccctga ccatccaggt tggcgccaac gacggtgaaa ctatcgatat cgatctgaag 540
cagatcaact ctcagaccct gggtctggac tcactgaacg tgcagaaagc gtatgatgtg 600
aaagatacag cagtaacaac gaaagcttat gccaataatg gtactacact ggatgtatcg 660
ggtcttgatg atgcagctat taaagcggct acgggtggta cgaatggtac ggcttctgta 720
accggtggtg cggttaaatt tgacgcagat aataacaagt actttgttac tattggtggc 780
tttactggtg ctgatgccgc caaaaatggc gattatgaag ttaacgttgc tactgacggt 840
acagtaaccc ttgcggctgg cgcaactaaa accacaatgc ctgctggtgc gacaactaaa 900
acagaagtac aggagttaaa agatacaccg gcagttgttt cagcagatgc taaaaatgcc 960
ttaattgctg gcggcgttga cgctaccgat gctaatggcg ctgagttggt caaaatgtct 1020
tataccgata aaaatggtaa gacaattgaa ggcggttatg cgcttaaagc tggcgataag 1080
tattacgccg cagattacga tgaagcgaca ggagcaatta aagctaaaac tacaagttat 1140
actgctgctg acggcactac caaaacagcg gctaaccaac tgggtggcgt agacggtaaa 1200
accgaagtcg ttactatcga cggtaaaacc tacaatgcca gcaaagccgc tggtcatgat 1260
ttcaaagcac aaccagagct ggcggaagca gccgctaaaa ccaccgaaaa cccgctgcag 1320
aaaattgatg ccgcgctggc gcaggtggat gcgctgcgct ctgatctggg tgcggtacaa 1380
aaccgtttca actctgctat caccaacctg ggcaataccg taaacaatct gtctgaagcg 1440
cgtagccgta tcgaagattc cgactacgcg accgaagttt ccaacatgtc tcgcgcgcag 1500
attctgcagc aggccggtac ttccgttctg gcgcaggcta accaggtccc gcagaacgtg 1560
ctgtctctgt tacgtaaggg caattcgaag cttgaaggtc aattggaatt ctctctgctg 1620
actgaagtag aaactccaac gcgtaatgaa tgggaatccc gttctagcga ctcctctgat 1680
cctctcgagt ccctgctgac ggaggttgaa accccgaccc gcaacgagtg ggaaagccgt 1740
tcctccgatt cctctgatcc ggagagcagc ctgctgaccg aggtagaaac cccgacccgt 1800
aatgagtggg aatctcgctc ctctgattct tctgacccgg gatcctctct gctgaccgaa 1860
gtggagactc cgactcgcaa cgaatgggag agccgttctt ctgactcctc tgacccgtct 1920
agataat 1927

<210> 583
<211> 749
<212> PRT
<213> Artificial Sequence

<220>
<223> pMT/STF2.4XM2e (H1 H5)

<400> 583
Met Lys Leu Cys Ile Leu Leu Ala Val Val Ala Phe Val Gly Leu Ser
1 5 10 15
Leu Gly Arg Ser Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu
20 25 30
Thr Gln Asn Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile
35 40 45
Glu Arg Leu Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala
50 55 60
Ala Gly Gln Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu
65 70 75 80
Thr Gln Ala Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr
85 90 95
Thr Glu Gly Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg
100 105 110
Glu Leu Ala Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu
115 120 125
Asp Ser Ile Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg
130 135 140
Val Ser Gly Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp
145 150 155 160
Asn Thr Leu Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp
165 170 175
Ile Asp Leu Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu

```
            180                 185                 190
Asn Val Gln Lys Ala Tyr Asp Val Lys Asp Thr Ala Val Thr Thr Lys
        195                 200                 205
Ala Tyr Ala Asn Asn Gly Thr Thr Leu Asp Val Ser Gly Leu Asp Asp
    210                 215                 220
Ala Ala Ile Lys Ala Ala Thr Gly Gly Thr Asn Gly Thr Ala Ser Val
225                 230                 235                 240
Thr Gly Gly Ala Val Lys Phe Asp Ala Asp Asn Asn Lys Tyr Phe Val
                245                 250                 255
Thr Ile Gly Gly Phe Thr Gly Ala Asp Ala Ala Lys Asn Gly Asp Tyr
            260                 265                 270
Glu Val Asn Val Ala Thr Asp Gly Thr Val Thr Leu Ala Ala Gly Ala
        275                 280                 285
Thr Lys Thr Thr Met Pro Ala Gly Ala Thr Thr Lys Thr Glu Val Gln
    290                 295                 300
Glu Leu Lys Asp Thr Pro Ala Val Val Ser Ala Asp Ala Lys Asn Ala
305                 310                 315                 320
Leu Ile Ala Gly Gly Val Asp Ala Thr Asp Ala Asn Gly Ala Glu Leu
                325                 330                 335
Val Lys Met Ser Tyr Thr Asp Lys Asn Gly Lys Thr Ile Glu Gly Gly
            340                 345                 350
Tyr Ala Leu Lys Ala Gly Asp Lys Tyr Tyr Ala Ala Asp Tyr Asp Glu
        355                 360                 365
Ala Thr Gly Ala Ile Lys Ala Lys Thr Thr Ser Tyr Thr Ala Ala Asp
    370                 375                 380
Gly Thr Thr Lys Thr Ala Ala Asn Gln Leu Gly Gly Val Asp Gly Lys
385                 390                 395                 400
Thr Glu Val Val Thr Ile Asp Gly Lys Thr Tyr Asn Ala Ser Lys Ala
                405                 410                 415
Ala Gly His Asp Phe Lys Ala Gln Pro Glu Leu Ala Glu Ala Ala Ala
            420                 425                 430
Lys Thr Thr Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala Gln
        435                 440                 445
Val Asp Ala Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe Asn
    450                 455                 460
Ser Ala Ile Thr Asn Leu Gly Asn Thr Val Asn Asn Leu Ser Glu Ala
465                 470                 475                 480
Arg Ser Arg Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn Met
                485                 490                 495
Ser Arg Ala Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala Gln
            500                 505                 510
Ala Asn Gln Val Pro Gln Asn Val Leu Ser Leu Leu Arg Lys Gly Asn
        515                 520                 525
Ser Lys Leu Glu Gly Gln Leu Glu Phe Ser Leu Leu Thr Glu Val Glu
    530                 535                 540
Thr Pro Ile Arg Asn Glu Trp Gly Ser Arg Ser Asn Asp Ser Ser Asp
545                 550                 555                 560
Pro Leu Glu Ser Leu Leu Thr Glu Val Glu Thr Pro Ile Arg Asn Glu
                565                 570                 575
Trp Gly Ser Arg Ser Asn Asp Ser Ser Asp Pro Gly Ser Ser Leu Leu
            580                 585                 590
Thr Glu Val Glu Thr Pro Ile Arg Asn Glu Trp Gly Ser Arg Ser Asn
        595                 600                 605
Asp Ser Ser Asp Pro Glu Leu Ser Leu Leu Thr Glu Val Glu Thr Pro
    610                 615                 620
Ile Arg Asn Glu Trp Gly Ser Arg Ser Asn Asp Ser Ser Asp Pro Ser
625                 630                 635                 640
Arg Ser Arg Glu Phe Ser Leu Leu Thr Glu Val Glu Thr Pro Thr Arg
                645                 650                 655
```

```
Asn Glu Trp Glu Ser Arg Ser Ser Asp Ser Ser Asp Pro Leu Glu Ser
            660                 665                 670
Leu Leu Thr Glu Val Glu Thr Pro Thr Arg Asn Glu Trp Glu Ser Arg
        675                 680                 685
Ser Ser Asp Ser Ser Asp Pro Glu Ser Ser Leu Leu Thr Glu Val Glu
    690                 695                 700
Thr Pro Thr Arg Asn Glu Trp Glu Ser Arg Ser Ser Asp Ser Ser Asp
705                 710                 715                 720
Pro Gly Ser Ser Leu Leu Thr Glu Val Glu Thr Pro Thr Arg Asn Glu
                725                 730                 735
Trp Glu Ser Arg Ser Ser Asp Ser Ser Asp Pro Ser Arg
            740                 745
```

<210> 584
<211> 2250
<212> DNA
<213> Artificial Sequence

<220>
<223> pMT/STF2.4XM2e (H1 H5)

<400> 584

```
atgaagttat gcatattact ggccgtcgtg gcctttgttg gcctctcgct cgggagatct 60
gcacaagtaa tcaacactaa cagtctgtcg ctgctgaccc agaataacct gaacaaatcc 120
cagtccgcac tgggcaccgc tatcgagcgt ctgtcttctg gtctgcgtat caacagcgcg 180
aaagacgatg cggcaggtca ggcgattgct aaccgtttca ccgcgaacat caaaggtctg 240
actcaggctt cccgtaacgc taacgacggt atctccattg cgcagaccac tgaaggcgcg 300
ctgaacgaaa tcaacaacaa cctgcagcgt gtgcgtgaac tggcggttca gtctgctaac 360
agcaccaact cccagtctga cctcgactcc atccaggctg aaatcaccca gcgcctgaac 420
gaaatcgacc gtgtatccgg ccagactcag ttcaacggcg tgaaagtcct ggcgcaggac 480
aacaccctga ccatccaggt tggcgccaac gacggtgaaa ctatcgatat cgatctgaag 540
cagatcaact ctcagaccct gggtctggac tcactgaacg tgcagaaagc gtatgatgtg 600
aaagatacag cagtaacaac gaaagcttat gccaataatg gtactacact ggatgtatcg 660
ggtcttgatg atgcagctat taaagcggct acgggtggta cgaatggtac ggcttctgta 720
accggtggtg cggttaaatt tgacgcagat aataacaagt actttgttac tattggtggc 780
tttactggtg ctgatgccgc caaaaatggc gattatgaag ttaacgttgc tactgacggt 840
acagtaaccc ttgcggctgg cgcaactaaa accacaatgc ctgctggtgc gacaactaaa 900
acagaagtac aggagttaaa agatacaccg gcagttgttt cagcagatgc taaaaatgcc 960
ttaattgctg gcggcgttga cgctaccgat gctaatggcg ctgagttggt caaaatgtct 1020
tataccgata aaaatggtaa gacaattgaa ggcggttatg cgcttaaagc tggcgataag 1080
tattacgccg cagattacga tgaagcgaca ggagcaatta aagctaaaac tacaagttat 1140
actgctgctg acggcactac caaaacagcg gctaaccaac tgggtggcgt agacggtaaa 1200
accgaagtcg ttactatcga cggtaaaacc tacaatgcca gcaaagccgc tggtcatgat 1260
ttcaaagcac aaccagagct ggcggaagca gccgctaaaa ccaccgaaaa cccgctgcag 1320
aaaattgatg ccgcgctggc gcaggtggat gcgctgcgct ctgatctggg tgcggtacaa 1380
aaccgtttca actctgctat caccaacctg ggcaataccg taaacaatct gtctgaagcg 1440
cgtagccgta tcgaagattc cgactacgcg accgaagttt ccaacatgtc tcgcgcgcag 1500
attctgcagc aggccggtac ttccgttctg gcgcaggcta accaggtccc gcagaacgtg 1560
ctgtctctgt tacgtaaggg caattcgaag cttgaaggtc aattggaatt cagcttgctg 1620
actgaggttg agaccccgat tcgcaacgaa tggggttccc gttccaacga ttcttccgac 1680
ccgctcgaga gcctgttgac cgaggttgaa acccctatcc gtaatgaatg gggctcccgt 1740
agcaacgact cttctgaccc aggatcctcc ctcttgaccg aagtggaaac gcctattcgt 1800
aatgagtggg gttctcgtag caatgacagc tcggacccgg agctctcgct gctgacggaa 1860
gtggagactc cgatccgtaa cgagtggggc tctcgctcta acgatagctc agacccgtct 1920
agatctagag aattctctct gctgactgaa gtagaaactc caacgcgtaa tgaatgggaa 1980
tcccgttcta gcgactcctc tgatcctctc gagtccctgc tgacggaggt tgaaaccccg 2040
acccgcaacg agtgggaaag ccgttcctcc gattcctctg atccggagag cagcctgctg 2100
```

accgaggtag aaaccccgac ccgtaatgag tgggaatctc gctcctctga ttcttctgac 2160
ccgggatcct ctctgctgac cgaagtggag actccgactc gcaacgaatg ggagagccgt 2220
tcttctgact cctctgaccc gtctagataa 2250

<210> 585
<211> 318
<212> PRT
<213> Artificial Sequence

<220>
<223> pMT/STF2delta

<400> 585

Met	Lys	Leu	Cys	Ile	Leu	Leu	Ala	Val	Val	Ala	Phe	Val	Gly	Leu	Ser
1				5					10					15	
Leu	Gly	Arg	Ser	Ala	Gln	Val	Ile	Asn	Thr	Asn	Ser	Leu	Ser	Leu	Leu
			20					25					30		
Thr	Gln	Asn	Asn	Leu	Asn	Lys	Ser	Gln	Ser	Ala	Leu	Gly	Thr	Ala	Ile
		35					40					45			
Glu	Arg	Leu	Ser	Ser	Gly	Leu	Arg	Ile	Asn	Ser	Ala	Lys	Asp	Asp	Ala
	50					55					60				
Ala	Gly	Gln	Ala	Ile	Ala	Asn	Arg	Phe	Thr	Ala	Asn	Ile	Lys	Gly	Leu
65					70					75					80
Thr	Gln	Ala	Ser	Arg	Asn	Ala	Asn	Asp	Gly	Ile	Ser	Ile	Ala	Gln	Thr
				85					90					95	
Thr	Glu	Gly	Ala	Leu	Asn	Glu	Ile	Asn	Asn	Asn	Leu	Gln	Arg	Val	Arg
			100					105					110		
Glu	Leu	Ala	Val	Gln	Ser	Ala	Asn	Ser	Thr	Asn	Ser	Gln	Ser	Asp	Leu
		115					120					125			
Asp	Ser	Ile	Gln	Ala	Glu	Ile	Thr	Gln	Arg	Leu	Asn	Glu	Ile	Asp	Arg
	130					135					140				
Val	Ser	Gly	Gln	Thr	Gln	Phe	Asn	Gly	Val	Lys	Val	Leu	Ala	Gln	Asp
145					150					155					160
Asn	Thr	Leu	Thr	Ile	Gln	Val	Gly	Ala	Asn	Asp	Gly	Glu	Thr	Ile	Asp
				165					170					175	
Ile	Asp	Leu	Lys	Gln	Ile	Asn	Ser	Gln	Thr	Leu	Gly	Leu	Asp	Ser	Leu
			180					185					190		
Asn	Val	His	Gly	Ala	Pro	Val	Asp	Pro	Ala	Ser	Pro	Trp	Thr	Glu	Asn
		195					200					205			
Pro	Leu	Gln	Lys	Ile	Asp	Ala	Ala	Leu	Ala	Gln	Val	Asp	Ala	Leu	Arg
	210					215					220				
Ser	Asp	Leu	Gly	Ala	Val	Gln	Asn	Arg	Phe	Asn	Ser	Ala	Ile	Thr	Asn
225					230					235					240
Leu	Gly	Asn	Thr	Val	Asn	Asn	Leu	Ser	Glu	Ala	Arg	Ser	Arg	Ile	Glu
				245					250					255	
Asp	Ser	Asp	Tyr	Ala	Thr	Glu	Val	Ser	Asn	Met	Ser	Arg	Ala	Gln	Ile
			260					265					270		
Leu	Gln	Gln	Ala	Gly	Thr	Ser	Val	Leu	Ala	Gln	Ala	Asn	Gln	Val	Pro
		275					280					285			
Gln	Asn	Val	Leu	Ser	Leu	Leu	Arg	Glu	Phe	Ser	Arg	Tyr	Pro	Ala	Gln
	290					295					300				
Trp	Arg	Pro	Leu	Thr	Arg	Thr	Gly	His	His	His	His	His	His		
305					310					315					

<210> 586
<211> 957
<212> DNA

<213> Artificial Sequence

<220>
<223> pMT/STF2delta

<400> 586
atgaagttat gcatattact ggccgtcgtg gcctttgttg gcctctcgct cgggagatct 60
gcacaagtaa tcaacactaa cagtctgtcg ctgctgaccc agaataacct gaacaaatcc 120
cagtccgcac tgggcaccgc tatcgagcgt ctgtcttctg gtctgcgtat caacagcgcg 180
aaagacgatg cggcaggtca ggcgattgct aaccgtttca ccgcgaacat caaaggtctg 240
actcaggctt cccgtaacgc taacgacggt atctccattg cgcagaccac tgaaggcgcg 300
ctgaacgaaa tcaacaacaa cctgcagcgt gtgcgtgaac tggcggttca gtctgctaac 360
agcaccaact cccagtctga cctcgactcc atccaggctg aaatcaccca gcgcctgaac 420
gaaatcgacc gtgtatccgg ccagactcag ttcaacggcg tgaaagtcct ggcgcaggac 480
aacaccctga ccatccaggt tggcgccaac gacggtgaaa ctatcgatat cgatctgaag 540
cagatcaact ctcagaccct gggtctggac tcactgaacg tgcatggagc gccggtggat 600
cctgctagcc catggaccga aaacccgctg cagaaaattg atgccgcgct ggcgcaggtg 660
gatgcgctgc gctctgatct gggtgcggta caaaaccgtt tcaactctgc tatcaccaac 720
ctgggcaata ccgtaaacaa tctgtctgaa gcgcgtagcc gtatcgaaga ttccgactac 780
gcgaccgaag tttccaacat gtctcgcgcg cagattttgc agcaggccgg tacttccgtt 840
ctggcgcagg ctaaccaggt cccgcagaac gtgctgtctc tgttacgtga attctctaga 900
tatccagcac agtggcggcc gctcacgcgt accggtcatc atcaccatca ccattga 957

<210> 587
<211> 402
<212> PRT
<213> Artificial Sequence

<220>
<223> pMT/STF2delta.4XM2e (H1)

<400> 587
Met Lys Leu Cys Ile Leu Leu Ala Val Val Ala Phe Val Gly Leu Ser
1 5 10 15
Leu Gly Arg Ser Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu
20 25 30
Thr Gln Asn Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile
35 40 45
Glu Arg Leu Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala
50 55 60
Ala Gly Gln Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu
65 70 75 80
Thr Gln Ala Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr
85 90 95
Thr Glu Gly Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg
100 105 110
Glu Leu Ala Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu
115 120 125
Asp Ser Ile Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg
130 135 140
Val Ser Gly Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp
145 150 155 160
Asn Thr Leu Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp
165 170 175
Ile Asp Leu Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu
180 185 190
Asn Val His Gly Ala Pro Val Asp Pro Ala Ser Pro Trp Thr Glu Asn

```
        195                 200                 205
Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala Leu Arg
    210                 215                 220
Ser Asp Leu Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile Thr Asn
225                 230                 235                 240
Leu Gly Asn Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg Ile Glu
                245                 250                 255
Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala Gln Ile
            260                 265                 270
Leu Gln Gln Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln Val Pro
        275                 280                 285
Gln Asn Val Leu Ser Leu Leu Arg Glu Phe Ser Leu Leu Thr Glu Val
    290                 295                 300
Glu Thr Pro Ile Arg Asn Glu Trp Gly Ser Arg Ser Asn Asp Ser Ser
305                 310                 315                 320
Asp Pro Leu Glu Ser Leu Leu Thr Glu Val Glu Thr Pro Ile Arg Asn
                325                 330                 335
Glu Trp Gly Ser Arg Ser Asn Asp Ser Ser Asp Pro Gly Ser Ser Leu
            340                 345                 350
Leu Thr Glu Val Glu Thr Pro Ile Arg Asn Glu Trp Gly Ser Arg Ser
        355                 360                 365
Asn Asp Ser Ser Asp Pro Glu Leu Ser Leu Leu Thr Glu Val Glu Thr
    370                 375                 380
Pro Ile Arg Asn Glu Trp Gly Ser Arg Ser Asn Asp Ser Ser Asp Pro
385                 390                 395                 400
Ser Arg
```

<210> 588
<211> 1203
<212> DNA
<213> Artificial Sequence

<220>
<223> pMT/STF2delta.4XM2e (H1)

<400> 588
atgtgcatat tactggccgt cgtggccttt gttggcctct cgctcgggag atctgcacaa 60
gtaatcaaca ctaacagtct gtcgctgctg acccagaata acctgaacaa atcccagtcc 120
gcactgggca ccgctatcga gcgtctgtct tctggtctgc gtatcaacag cgcgaaagac 180
gatgcggcag gtcaggcgat tgctaaccgt ttcaccgcga acatcaaagg tctgactcag 240
gcttcccgta acgctaacga cggtatctcc attgcgcaga ccactgaagg cgcgctgaac 300
gaaatcaaca acaacctgca gcgtgtgcgt gaactggcgg ttcagtctgc taacagcacc 360
aactcccagt ctgacctcga ctccatccag gctgaaatca cccagcgcct gaacgaaatc 420
gaccgtgtat ccggccagac tcagttcaac ggcgtgaaag tcctggcgca ggacaacacc 480
ctgaccatcc aggttggcgc caacgacggt gaaactatcg atatcgatct gaagcagatc 540
aactctcaga ccctgggtct ggactcactg aacgtgcatg gagcgccggt ggatcctgct 600
agcccatgga ccgaaaaccc gctgcagaaa attgatgccg cgctggcgca ggtggatgcg 660
ctgcgctctg atctgggtgc ggtacaaaac cgtttcaact ctgctatcac caacctgggc 720
aataccgtaa acaatctgtc tgaagcgcgt agccgtatcg aagattccga ctacgcgacc 780
gaagtttcca acatgtctcg cgcgcagatt ttgcagcagg ccggtacttc cgttctggcg 840
caggctaacc aggtcccgca gaacgtgctg tctctgttac gtgaattcag cttgctgact 900
gaggttgaga ccccgattcg caacgaatgg ggttcccgtt ccaacgattc ttccgacccg 960
ctcgagagcc tgttgaccga ggttgaaacc cctatccgta atgaatgggg ctcccgtagc 1020
aacgactctt ctgacccagg atcctccctc ttgaccgaag tggaaacgcc tattcgtaat 1080
gagtggggtt ctcgtagcaa tgacagctcg gacccggagc tctcgctgct gacggaagtg 1140
gagactccga tccgtaacga gtggggctct cgctctaacg atagctcaga cccgtctaga 1200

taa 1203

<210> 589
<211> 402
<212> PRT
<213> Artificial Sequence

<220>
<223> pMT/STF2delta.4XM2e (H5)

<400> 589

```
Met Lys Leu Cys Ile Leu Leu Ala Val Val Ala Phe Val Gly Leu Ser
 1               5                  10                  15
Leu Gly Arg Ser Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu
            20                  25                  30
Thr Gln Asn Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile
        35                  40                  45
Glu Arg Leu Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala
    50                  55                  60
Ala Gly Gln Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu
65                  70                  75                  80
Thr Gln Ala Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr
                85                  90                  95
Thr Glu Gly Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg
            100                 105                 110
Glu Leu Ala Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu
        115                 120                 125
Asp Ser Ile Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg
    130                 135                 140
Val Ser Gly Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp
145                 150                 155                 160
Asn Thr Leu Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp
                165                 170                 175
Ile Asp Leu Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu
            180                 185                 190
Asn Val His Gly Ala Pro Val Asp Pro Ala Ser Pro Trp Thr Glu Asn
        195                 200                 205
Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala Leu Arg
    210                 215                 220
Ser Asp Leu Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile Thr Asn
225                 230                 235                 240
Leu Gly Asn Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg Ile Glu
                245                 250                 255
Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala Gln Ile
            260                 265                 270
Leu Gln Gln Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln Val Pro
        275                 280                 285
Gln Asn Val Leu Ser Leu Leu Arg Glu Phe Ser Leu Leu Thr Glu Val
    290                 295                 300
Glu Thr Pro Thr Arg Asn Glu Trp Glu Ser Arg Ser Ser Asp Ser Ser
305                 310                 315                 320
Asp Pro Leu Glu Ser Leu Leu Thr Glu Val Glu Thr Pro Thr Arg Asn
                325                 330                 335
Glu Trp Glu Ser Arg Ser Ser Asp Ser Ser Asp Pro Glu Ser Ser Leu
            340                 345                 350
Leu Thr Glu Val Glu Thr Pro Thr Arg Asn Glu Trp Glu Ser Arg Ser
        355                 360                 365
Ser Asp Ser Ser Asp Pro Gly Ser Ser Leu Leu Thr Glu Val Glu Thr
```

```
    370                 375                 380
Pro Thr Arg Asn Glu Trp Glu Ser Arg Ser Ser Asp Ser Ser Asp Pro
385                 390                 395                 400
Ser Arg
```

<210> 590
<211> 1206
<212> DNA
<213> Artificial Sequence

<220>
<223> pMT/STF2delta.4XM2e (H5)

<400> 590
atgtgcatat tactggccgt cgtggccttt gttggcctct cgctcgggag atctgcacaa 60
gtaatcaaca ctaacagtct gtcgctgctg acccagaata acctgaacaa atcccagtcc 120
gcactgggca ccgctatcga gcgtctgtct tctggtctgc gtatcaacag cgcgaaagac 180
gatgcggcag gtcaggcgat tgctaaccgt ttcaccgcga acatcaaagg tctgactcag 240
gcttcccgta acgctaacga cggtatctcc attgcgcaga ccactgaagg cgcgctgaac 300
gaaatcaaca acaacctgca gcgtgtgcgt gaactggcgg ttcagtctgc taacagcacc 360
aactcccagt ctgacctcga ctccatccag gctgaaatca cccagcgcct gaacgaaatc 420
gaccgtgtat ccggccagac tcagttcaac ggcgtgaaag tcctggcgca ggacaacacc 480
ctgaccatcc aggttggcgc caacgacggt gaaactatcg atatcgatct gaagcagatc 540
aactctcaga ccctgggtct ggactcactg aacgtgcatg gagcgccggt ggatcctgct 600
agcccatgga ccgaaaaccc gctgcagaaa attgatgccg cgctggcgca ggtggatgcg 660
ctgcgctctg atctgggtgc ggtacaaaac cgtttcaact ctgctatcac caacctgggc 720
aataccgtaa acaatctgtc tgaagcgcgt agccgtatcg aagattccga ctacgcgacc 780
gaagtttcca acatgtctcg cgcgcagatt ttgcagcagg ccggtacttc cgttctggcg 840
caggctaacc aggtcccgca gaacgtgctg tctctgttac gtgaattctc tctgctgact 900
gaagtagaaa ctccaacgcg taatgaatgg gaatcccgtt ctagcgactc ctctgatcct 960
ctcgagtccc tgctgacgga ggttgaaacc ccgacccgca acgagtggga aagccgttcc 1020
tccgattcct ctgatccgga gagcagcctg ctgaccgagg tagaaacccc gacccgtaat 1080
gagtgggaat ctcgctcctc tgattcttct gacccgggat cctctctgct gaccgaagtg 1140
gagactccga ctcgcaacga atgggagagc cgttcttctg actcctctga cccgtctaga 1200
taataa 1206

<210> 591
<211> 508
<212> PRT
<213> Artificial Sequence

<220>
<223> pMT/STF2delta.4XM2e (H1H5)

<400> 591
```
Met Lys Leu Cys Ile Leu Leu Ala Val Val Ala Phe Val Gly Leu Ser
 1               5                  10                  15
Leu Gly Arg Ser Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu
            20                  25                  30
Thr Gln Asn Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile
        35                  40                  45
Glu Arg Leu Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala
    50                  55                  60
Ala Gly Gln Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu
65                  70                  75                  80
```

```
Thr Gln Ala Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr
                85                  90                  95
Thr Glu Gly Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg
            100                 105                 110
Glu Leu Ala Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu
        115                 120                 125
Asp Ser Ile Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg
    130                 135                 140
Val Ser Gly Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp
145                 150                 155                 160
Asn Thr Leu Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp
                165                 170                 175
Ile Asp Leu Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu
            180                 185                 190
Asn Val His Gly Ala Pro Val Asp Pro Ala Ser Pro Trp Thr Glu Asn
        195                 200                 205
Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala Leu Arg
    210                 215                 220
Ser Asp Leu Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile Thr Asn
225                 230                 235                 240
Leu Gly Asn Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg Ile Glu
                245                 250                 255
Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala Gln Ile
            260                 265                 270
Leu Gln Gln Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln Val Pro
        275                 280                 285
Gln Asn Val Leu Ser Leu Leu Arg Glu Phe Ser Leu Leu Thr Glu Val
    290                 295                 300
Glu Thr Pro Ile Arg Asn Glu Trp Gly Ser Arg Ser Asn Asp Ser Ser
305                 310                 315                 320
Asp Pro Leu Glu Ser Leu Leu Thr Glu Val Glu Thr Pro Ile Arg Asn
                325                 330                 335
Glu Trp Gly Ser Arg Ser Asn Asp Ser Ser Asp Pro Gly Ser Ser Leu
            340                 345                 350
Leu Thr Glu Val Glu Thr Pro Ile Arg Asn Glu Trp Gly Ser Arg Ser
        355                 360                 365
Asn Asp Ser Ser Asp Pro Glu Leu Ser Leu Leu Thr Glu Val Glu Thr
    370                 375                 380
Pro Ile Arg Asn Glu Trp Gly Ser Arg Ser Asn Asp Ser Ser Asp Pro
385                 390                 395                 400
Ser Arg Gln Phe Ser Leu Leu Thr Glu Val Glu Thr Pro Thr Arg Asn
                405                 410                 415
Glu Trp Glu Ser Arg Ser Ser Asp Ser Ser Asp Pro Leu Glu Ser Leu
            420                 425                 430
Leu Thr Glu Val Glu Thr Pro Thr Arg Asn Glu Trp Glu Ser Arg Ser
        435                 440                 445
Ser Asp Ser Ser Asp Pro Glu Ser Ser Leu Leu Thr Glu Val Glu Thr
    450                 455                 460
Pro Thr Arg Asn Glu Trp Glu Ser Arg Ser Ser Asp Ser Ser Asp Pro
465                 470                 475                 480
Gly Ser Ser Leu Leu Thr Glu Val Glu Thr Pro Thr Arg Asn Glu Trp
                485                 490                 495
Glu Ser Arg Ser Ser Asp Ser Ser Asp Pro Ser Arg
            500                 505
```

<210> 592
<211> 1527
<212> DNA

<213> Artificial Sequence

<220>
<223> pMT/STF2delta.4XM2e (H1H5)

<400> 592
atgtgcatat tactggccgt cgtggccttt gttggcctct cgctcgggag atctgcacaa 60
gtaatcaaca ctaacagtct gtcgctgctg acccagaata acctgaacaa atcccagtcc 120
gcactgggca ccgctatcga gcgtctgtct tctggtctgc gtatcaacag cgcgaaagac 180
gatgcggcag gtcaggcgat tgctaaccgt ttcaccgcga acatcaaagg tctgactcag 240
gcttcccgta acgctaacga cggtatctcc attgcgcaga ccactgaagg cgcgctgaac 300
gaaatcaaca acaacctgca gcgtgtgcgt gaactggcgg ttcagtctgc taacagcacc 360
aactcccagt ctgacctcga ctccatccag gctgaaatca cccagcgcct gaacgaaatc 420
gaccgtgtat ccggccagac tcagttcaac ggcgtgaaag tcctggcgca ggacaacacc 480
ctgaccatcc aggttggcgc caacgacggt gaaactatcg atatcgatct gaagcagatc 540
aactctcaga ccctgggtct ggactcactg aacgtgcatg gagcgccggt ggatcctgct 600
agcccatgga ccgaaaaccc gctgcagaaa attgatgccg cgctggcgca ggtggatgcg 660
ctgcgctctg atctgggtgc ggtacaaaac cgtttcaact ctgctatcac caacctgggc 720
aataccgtaa acaatctgtc tgaagcgcgt agccgtatcg aagattccga ctacgcgacc 780
gaagtttcca acatgtctcg cgcgcagatt ttgcagcagg ccggtacttc cgttctggcg 840
caggctaacc aggtcccgca gaacgtgctg tctctgttac gtgaattcag cttgctgact 900
gaggttgaga ccccgattcg caacgaatgg ggttcccgtt ccaacgattc ttccgacccg 960
ctcgagagcc tgttgaccga ggttgaaacc cctatccgta atgaatgggg ctcccgtagc 1020
aacgactctt ctgacccagg atcctccctc ttgaccgaag tggaaacgcc tattcgtaat 1080
gagtggggtt ctcgtagcaa tgacagctcg gacccggagc tctcgctgct gacggaagtg 1140
gagactccga tccgtaacga gtggggctct cgctctaacg atagctcaga cccgtctaga 1200
tctagagaat tctctctgct gactgaagta gaaactccaa cgcgtaatga atgggaatcc 1260
cgttctagcg actcctctga tcctctcgag tccctgctga cggaggttga aaccccgacc 1320
cgcaacgagt gggaaagccg ttcctccgat tcctctgatc cggagagcag cctgctgacc 1380
gaggtagaaa ccccgacccg taatgagtgg gaatctcgct cctctgattc ttctgacccg 1440
ggatcctctc tgctgaccga agtggagact ccgactcgca acgaatggga gagccgttct 1500
tctgactcct ctgacccgtc tagataa 1527

<210> 593
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> M2

<400> 593
gaggttgaga ccccgattcg caacgaatgg ggt 33

<210> 594
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> Linker

<400> 594
Gly Ala Pro Val Asp Pro Ala Ser Pro Trp
1 5 10

<210> 595
<211> 277
<212> PRT
<213> Artificial Sequence

<220>
<223> S. muenchen

<400> 595
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
1 5 10 15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
20 25 30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
35 40 45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
50 55 60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65 70 75 80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
85 90 95
Val Gln Ser Ala Asn Gly Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
100 105 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
115 120 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
130 135 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145 150 155 160
Lys Glu Ile Ser Ser Lys Thr Leu Asp Lys His Asn Phe Arg Thr Gly
165 170 175
Gly Glu Leu Lys Glu Val Asn Thr Asp Lys Thr Glu Asn Pro Leu Gln
180 185 190
Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Thr Leu Arg Ser Asp Leu
195 200 205
Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile Thr Asn Leu Gly Asn
210 215 220
Thr Val Asn Asn Leu Ser Ser Ala Arg Ser Arg Ile Glu Asp Ser Asp
225 230 235 240
Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala Gln Ile Leu Gln Gln
245 250 255
Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln Val Pro Gln Asn Val
260 265 270
Leu Ser Leu Leu Arg
275

<210> 596
<211> 789
<212> DNA
<213> Artificial Sequence

<220>
<223> S. muenchen

<400> 596

atggcacaag tcattaatac aaacagcctg tcgctgttga cccagaataa cctgaacaaa 60
tcccagtccg ctctgggcac cgctatcgag cgtctgtctt ccggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacggtacta actcccagtc tgaccttgac tctatccagg ctgaaatcac ccagcgtctg 360
aacgaaatcg accgtgtatc cggtcagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggtgcc aacgacggtg aaactattga tattgattta 480
aaagaaatta gctctaaaac actgacagat aagactgaaa acccactgca gaaaattgat 540
gctgccttgg cacaggttga tacacttcgt tctgacctgg gtgcggtaca gaaccgtttc 600
aactccgcta tcaccaacct gggcaatacc gtaaataacc tgtcttctgc ccgtagccgt 660
atcgaagatt ccgactacgc gaccgaagtc tccaacatgt ctcgcgcgca gattctgcag 720
caggccggta cctccgttct ggcgcaggct aaccaggttc cgcaaaacgt cctctcttta 780
ctgcgttaa 789

<210> 597
<211> 102
<212> PRT
<213> Artificial Sequence

<220>
<223> H1 and H5

<400> 597
Ser Leu Leu Thr Glu Val Glu Thr Pro Thr Arg Asn Glu Trp Glu Ser
1 5 10 15
Arg Ser Ser Asp Ser Ser Asp Pro Leu Glu Ser Leu Leu Thr Glu Val
20 25 30
Glu Thr Pro Thr Arg Asn Glu Trp Glu Ser Arg Ser Ser Asp Ser Ser
35 40 45
Asp Pro Glu Ser Ser Leu Leu Thr Glu Val Glu Thr Pro Thr Arg Asn
50 55 60
Glu Trp Glu Ser Arg Ser Ser Asp Ser Ser Asp Pro Gly Ser Ser Leu
65 70 75 80
Leu Thr Glu Val Glu Thr Pro Thr Arg Asn Glu Trp Glu Ser Arg Ser
85 90 95
Ser Asp Ser Ser Asp Pro
100

<210> 598
<211> 306
<212> DNA
<213> Artificial Sequence

<220>
<223> H1 and H5

<400> 598
tctctgctga ctgaagtaga aactccaacg cgtaatgaat gggaatcccg ttctagcgac 60
tcctctgatc ctctcgagtc cctgctgacg gaggttgaaa ccccgacccg caacgagtgg 120
gaaagccgtt cctccgattc ctctgatccg gagagcagcc tgctgaccga ggtagaaacc 180
ccgacccgta atgagtggga atctcgctcc tctgattctt ctgacccggg atcctctctg 240
ctgaccgaag tggagactcc gactcgcaac gaatgggaga gccgttcttc tgactcctct 300
gacccg 306

<210> 599

<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> HA

<400> 599
Pro Ala Lys Leu Leu Lys Glu Arg Gly Phe Phe Gly Ala Ile Ala Gly
1 5 10 15
Phe Leu Glu Glu
20

<210> 600
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> HA

<400> 600
Asn Ile Pro Ser Ile Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly
1 5 10 15
Phe Ile Glu Glu
20

<210> 601
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> HA

<400> 601
Asn Val Pro Glu Lys Gln Thr Arg Gly Ile Phe Gly Ala Ile Ala Gly
1 5 10 15
Phe Ile Glu Glu
20

<210> 602
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> HA

<400> 602
Arg Glu Arg Arg Arg Lys Lys Arg Gly Leu Phe Gly Ala Ile Ala Gly
1 5 10 15

Phe Ile Glu Glu
20

<210> 603
<211> 72
<212> DNA
<213> Artificial Sequence

<220>
<223> M2

<400> 603
agcttgctga ctgaggttga gaccccgatt cgcaacgaat ggggttcccg ttccaacgat 60
tcttccgacc cg 72

<210> 604
<211> 506
<212> PRT
<213> S. typhimurium

<400> 604
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
1 5 10 15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
20 25 30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
35 40 45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
50 55 60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65 70 75 80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
85 90 95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
100 105 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
115 120 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
130 135 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145 150 155 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val Gln
165 170 175
Lys Ala Tyr Asp Val Lys Asp Thr Ala Val Thr Thr Lys Ala Tyr Ala
180 185 190
Asn Asn Gly Thr Thr Leu Asp Val Ser Gly Leu Asp Asp Ala Ala Ile
195 200 205
Lys Ala Ala Thr Gly Gly Thr Asn Gly Thr Ala Ser Val Thr Gly Gly
210 215 220
Ala Val Lys Phe Asp Ala Asp Asn Asn Lys Tyr Phe Val Thr Ile Gly
225 230 235 240
Gly Phe Thr Gly Ala Asp Ala Ala Lys Asn Gly Asp Tyr Glu Val Asn
245 250 255
Val Ala Thr Asp Gly Thr Val Thr Leu Ala Ala Gly Ala Thr Lys Thr
260 265 270
Thr Met Pro Ala Gly Ala Thr Thr Lys Thr Glu Val Gln Glu Leu Lys
275 280 285
Asp Thr Pro Ala Val Val Ser Ala Asp Ala Lys Asn Ala Leu Ile Ala

```
    290                 295                 300
Gly Gly Val Asp Ala Thr Asp Ala Asn Gly Ala Glu Leu Val Lys Met
305                 310                 315                 320
Ser Tyr Thr Asp Lys Asn Gly Lys Thr Ile Glu Gly Gly Tyr Ala Leu
                325                 330                 335
Lys Ala Gly Asp Lys Tyr Tyr Ala Ala Asp Tyr Asp Glu Ala Thr Gly
            340                 345                 350
Ala Ile Lys Ala Lys Thr Thr Ser Tyr Thr Ala Ala Asp Gly Thr Thr
        355                 360                 365
Lys Thr Ala Ala Asn Gln Leu Gly Gly Val Asp Gly Lys Thr Glu Val
    370                 375                 380
Val Thr Ile Asp Gly Lys Thr Tyr Asn Ala Ser Lys Ala Ala Gly His
385                 390                 395                 400
Asp Phe Lys Ala Gln Pro Glu Leu Ala Glu Ala Ala Ala Lys Thr Thr
                405                 410                 415
Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala
            420                 425                 430
Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile
        435                 440                 445
Thr Asn Leu Gly Asn Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg
    450                 455                 460
Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala
465                 470                 475                 480
Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln
                485                 490                 495
Val Pro Gln Asn Val Leu Ser Leu Leu Arg
            500                 505
```

<210> 605
<211> 1517
<212> DNA
<213> S. typhimurium

<400> 605

```
tggcacaagt aatcaacact aacagtctgt cgctgctgac ccagaataac ctgaacaaat 60
cccagtccgc actgggcacc gctatcgagc gtctgtcttc tggtctgcgt atcaacagcg 120
cgaaagacga tgcggcaggt caggcgattg ctaaccgttt caccgcgaac atcaaaggtc 180
tgactcaggc ttcccgtaac gctaacgacg gtatctccat tgcgcagacc actgaaggcg 240
cgctgaacga aatcaacaac aacctgcagc gtgtgcgtga actggcggtt cagtctgcta 300
acagcaccaa ctcccagtct gacctcgact ccatccaggc tgaaatcacc cagcgcctga 360
acgaaatcga ccgtgtatcc ggccagactc agttcaacgg cgtgaaagtc ctggcgcagg 420
acaacaccct gaccatccag gttggcgcca acgacggtga aactatcgat atcgatctga 480
agcagatcaa ctctcagacc ctgggtctgg actcactgaa cgtgcagaaa gcgtatgatg 540
tgaaagatac agcagtaaca acgaaagctt atgccaataa tggtactaca ctggatgtat 600
cgggtcttga tgatgcagct attaaagcgg ctacgggtgg tacgaatggt acggcttctg 660
taaccggtgg tgcggttaaa tttgacgcag ataataacaa gtactttgtt actattggtg 720
gctttactgg tgctgatgcc gccaaaaatg gcgattatga agttaacgtt gctactgacg 780
gtacagtaac ccttgcggct ggcgcaacta aaaccacaat gcctgctggt gcgacaacta 840
aaacagaagt acaggagtta aaagatacac cggcagttgt ttcagcagat gctaaaaatg 900
ccttaattgc tggcggcgtt gacgctaccg atgctaatgg cgctgagttg gtcaaaatgt 960
cttataccga taaaaatggt aagacaattg aaggcggtta tgcgcttaaa gctggcgata 1020
agtattacgc cgcagattac gatgaagcga caggagcaat taaagctaaa actacaagtt 1080
atactgctgc tgacggcact accaaaacag cggctaacca actgggtggc gtagacggta 1140
aaaccgaagt cgttactatc gacggtaaaa cctacaatgc cagcaaagcc gctggtcatg 1200
atttcaaagc acaaccagag ctggcggaag cagccgctaa aaccaccgaa aacccgctgc 1260
agaaaattga tgccgcgctg gcgcaggtgg atgcgctgcg ctctgatctg ggtgcggtac 1320
aaaaccgttt caactctgct atcaccaacc tgggcaatac cgtaaacaat ctgtctgaag 1380
cgcgtagccg tatcgaagat tccgactacg cgaccgaagt ttccaacatg tctcgcgcgc 1440
```

agattctgca gcaggccggt acttccgttc tggcgcaggc taaccaggtc ccgcagaacg 1500
tgctgtctct gttacgt 1517

<210> 606
<211> 277
<212> PRT
<213> S. typhimurium

<400> 606
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
1 5 10 15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
20 25 30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
35 40 45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
50 55 60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65 70 75 80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
85 90 95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
100 105 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
115 120 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
130 135 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145 150 155 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val His
165 170 175
Gly Ala Pro Val Asp Pro Ala Ser Pro Trp Thr Glu Asn Pro Leu Gln
180 185 190
Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala Leu Arg Ser Asp Leu
195 200 205
Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile Thr Asn Leu Gly Asn
210 215 220
Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg Ile Glu Asp Ser Asp
225 230 235 240
Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala Gln Ile Leu Gln Gln
245 250 255
Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln Val Pro Gln Asn Val
260 265 270
Leu Ser Leu Leu Arg
275

<210> 607
<211> 832
<212> DNA
<213> S. typhimurium

<400> 607
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360

aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcatgg agcgccggtg 540
gatcctgcta gcccatggac cgaaaacccg ctgcagaaaa ttgatgccgc gctggcgcag 600
gtggatgcgc tgcgctctga tctgggtgcg gtacaaaacc gtttcaactc tgctatcacc 660
aacctgggca ataccgtaaa caatctgtct gaagcgcgta gccgtatcga agattccgac 720
tacgcgaccg aagtttccaa catgtctcgc gcgcagattt tgcagcaggc cggtacttcc 780
gttctggcgc aggctaacca ggtcccgcag aacgtgctgt ctctgttacg tg 832

<210> 608
<211> 1215
<212> DNA
<213> Artificial Sequence

<220>
<223> pET/STF2delta.JEIII+

<400> 608
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcatgg agcgccggtg 540
gatcctgcta gcccatggac cgaaaacccg ctgcagaaaa ttgatgccgc gctggcgcag 600
gtggatgcgc tgcgctctga tctgggtgcg gtacaaaacc gtttcaactc tgctatcacc 660
aacctgggca ataccgtaaa caatctgtct gaagcgcgta gccgtatcga agattccgac 720
tacgcgaccg aagtttccaa catgtctcgc gcgcagattt tgcagcaggc cggtacttcc 780
gttctggcgc aggctaacca ggtcccgcag aacgtgctgt ctctgttacg tgaattctgc 840
agatatccag cacagtggcg gccgctcatg gacaaactgg ctctgaaagg cacaacctat 900
ggcatgtgta cagaaaaatt ctcgttcgcg aaaaatccgg tggacactgg tcacggaaca 960
gttgtcattg aactctccta ctctgggagt gatggcccct gcaaaattcc gattgtttcc 1020
gttgcgagcc tcaatgacat gacccccgtt gggcggctgg tgacagtgaa ccccttcgtc 1080
gcgacttcca gtgccaactc aaaggtgctg gtcgagatgg aacccccctt cggagactcc 1140
tacatcgtag ttggaagggg agacaagcag atcaaccacc attggcacaa agctggaagc 1200
acgctgggca aggcc 1215

<210> 609
<211> 405
<212> PRT
<213> Artificial Sequence

<220>
<223> pET/STF2delta.JEIII+

<400> 609
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
 1   5   10   15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
            20   25   30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
        35   40   45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
    50   55   60

```
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65                  70                  75                  80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
                85                  90                  95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
            100                 105                 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
        115                 120                 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
    130                 135                 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145                 150                 155                 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val His
                165                 170                 175
Gly Ala Pro Val Asp Pro Ala Ser Pro Trp Thr Glu Asn Pro Leu Gln
            180                 185                 190
Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala Leu Arg Ser Asp Leu
        195                 200                 205
Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile Thr Asn Leu Gly Asn
    210                 215                 220
Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg Ile Glu Asp Ser Asp
225                 230                 235                 240
Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala Gln Ile Leu Gln Gln
                245                 250                 255
Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln Val Pro Gln Asn Val
            260                 265                 270
Leu Ser Leu Leu Arg Glu Phe Cys Arg Tyr Pro Ala Gln Trp Arg Pro
        275                 280                 285
Leu Met Asp Lys Leu Ala Leu Lys Gly Thr Thr Tyr Gly Met Cys Thr
    290                 295                 300
Glu Lys Phe Ser Phe Ala Lys Asn Pro Val Asp Thr Gly His Gly Thr
305                 310                 315                 320
Val Val Ile Glu Leu Ser Tyr Ser Gly Ser Asp Gly Pro Cys Lys Ile
                325                 330                 335
Pro Ile Val Ser Val Ala Ser Leu Asn Asp Met Thr Pro Val Gly Arg
            340                 345                 350
Leu Val Thr Val Asn Pro Phe Val Ala Thr Ser Ser Ala Asn Ser Lys
        355                 360                 365
Val Leu Val Glu Met Glu Pro Pro Phe Gly Asp Ser Tyr Ile Val Val
    370                 375                 380
Gly Arg Gly Asp Lys Gln Ile Asn His His Trp His Lys Ala Gly Ser
385                 390                 395                 400
Thr Leu Gly Lys Ala
                405
```

<210> 610
<211> 115
<212> PRT
<213> Artificial Sequence

<220>
<223> West Nile domain III

<400> 610

```
Met Glu Lys Leu Gln Leu Lys Gly Thr Thr Tyr Gly Val Cys Ser Lys
 1               5                  10                  15
Ala Phe Lys Phe Leu Gly Thr Pro Ala Asp Thr Gly His Gly Thr Val
```

```
            20                  25                  30
Val Leu Glu Leu Gln Tyr Thr Gly Thr Asp Gly Pro Cys Lys Val Pro
        35                  40                  45
Ile Ser Ser Val Ala Ser Leu Asn Asp Leu Thr Pro Val Gly Arg Leu
    50                  55                  60
Val Thr Val Asn Pro Phe Val Ser Val Ala Thr Ala Asn Ala Lys Val
65                  70                  75                  80
Leu Ile Glu Leu Glu Pro Pro Phe Gly Asp Ser Tyr Ile Val Val Gly
                85                  90                  95
Arg Gly Glu Gln Gln Ile Asn His His Trp His Lys Ser Gly Ser Ser
            100                 105                 110
Ile Gly Lys
        115


<210> 611
<211> 103
<212> PRT
<213> Artificial Sequence

<220>
<223> West Nile domain III


<400> 611
Gly Thr Thr Tyr Gly Val Cys Ser Lys Ala Phe Lys Phe Ala Arg Thr
 1                  5                  10                  15
Pro Ala Asp Thr Gly His Gly Thr Val Val Leu Glu Leu Gln Tyr Thr
            20                  25                  30
Gly Lys Asp Gly Pro Cys Lys Val Pro Ile Ser Ser Val Ala Ser Leu
        35                  40                  45
Asn Asp Leu Thr Pro Val Gly Arg Leu Val Thr Val Asn Pro Phe Val
    50                  55                  60
Ser Val Ala Thr Ala Asn Ser Lys Val Leu Ile Glu Leu Glu Pro Pro
65                  70                  75                  80
Phe Gly Asp Ser Tyr Ile Val Val Gly Arg Gly Glu Gln Gln Ile Asn
                85                  90                  95
His His Trp His Lys Ser Gly
            100


<210> 612
<211> 102
<212> PRT
<213> Artificial Sequence

<220>
<223> West Nile domain III


<400> 612
Gly Thr Thr Tyr Gly Val Cys Ser Lys Ala Phe Lys Phe Leu Gly Thr
 1                  5                  10                  15
Pro Ala Asp Thr Gly His Gly Thr Val Val Leu Glu Leu Gln Tyr Thr
            20                  25                  30
Gly Thr Asp Gly Pro Cys Lys Val Pro Ile Ser Ser Val Ala Ser Leu
        35                  40                  45
Asn Asp Leu Thr Pro Val Gly Arg Leu Val Thr Val Asn Pro Phe Val
    50                  55                  60
```

```
Ser Val Ala Thr Ala Asn Ala Lys Val Leu Ile Glu Leu Glu Pro Pro
65                  70                  75                  80
Phe Gly Asp Ser Tyr Val Val Gly Arg Gly Glu Gln Gln Ile Asn His
                85                  90                  95
His Trp His Lys Ser Gly
            100


<210> 613
<211> 27
<212> PRT
<213> Artificial Sequence

<220>
<223> West Nile domain III


<400> 613
Glu Leu Glu Pro Pro Phe Gly Asp Ser Tyr Ile Val Val Gly Arg Gly
 1               5                  10                  15
Glu Gln Gln Ile Asn His His Trp His Lys Ser
            20                  25


<210> 614
<211> 345
<212> DNA
<213> Artificial Sequence

<220>
<223> West Nile domain III


<400> 614
atggaaaaat tgcagttgaa gggaacaacc tatggcgtct gttcaaaggc tttcaagttt 60
cttgggactc ccgcagacac aggtcacggc actgtggtgt tggaattgca gtacactggc 120
acggatggac cttgcaaagt tcctatctcg tcagtggctt cattgaacga cctaacgcca 180
gtgggcagat tggtcactgt caacccttttt gtttcagtgg ccacggccaa cgctaaggtc 240
ctgattgaat tggaaccacc ctttggagac tcatacatag tggtgggcag aggagaacaa 300
cagatcaatc accattggca caagtctgga agcagcattg gcaaa                  345

<210> 615
<211> 96
<212> PRT
<213> Artificial Sequence

<220>
<223> Langat virus


<400> 615
Gly Leu Thr Tyr Thr Val Cys Asp Lys Thr Lys Phe Thr Trp Lys Arg
 1               5                  10                  15
Ala Pro Thr Asp Ser Gly His Asp Thr Val Val Met Glu Val Gly Phe
            20                  25                  30
Ser Gly Thr Arg Pro Cys Arg Ile Pro Val Arg Ala Val Ala His Gly
        35                  40                  45
Val Pro Glu Val Asn Val Ala Met Leu Ile Thr Pro Asn Pro Thr Met
    50                  55                  60
```

```
Glu Asn Asn Gly Gly Gly Phe Ile Glu Met Gln Leu Pro Pro Gly Asp
65                  70                  75                  80
Asn Ile Ile Tyr Val Gly Asp Leu Asn His Gln Trp Phe Gln Lys Gly
                85                  90                  95


<210> 616
<211> 103
<212> PRT
<213> Artificial Sequence

<220>
<223> Kunjin virus


<400> 616
Gly Thr Thr Tyr Gly Val Cys Ser Lys Ala Phe Arg Phe Leu Gly Thr
 1               5                  10                  15
Pro Ala Asp Thr Gly His Gly Thr Val Val Leu Glu Leu Gln Tyr Thr
            20                  25                  30
Gly Thr Asp Gly Pro Cys Lys Ile Pro Ile Ser Ser Val Ala Ser Leu
        35                  40                  45
Asn Asp Leu Thr Pro Val Gly Arg Leu Val Thr Val Asn Pro Phe Val
    50                  55                  60
Ser Val Ser Thr Ala Asn Ala Lys Val Leu Ile Glu Leu Glu Pro Pro
65                  70                  75                  80
Phe Gly Asp Ser Tyr Ile Val Val Gly Arg Gly Glu Gln Gln Ile Asn
                85                  90                  95
His His Trp His Lys Ser Gly
            100


<210> 617
<211> 103
<212> PRT
<213> Artificial Sequence

<220>
<223> Murray Valley virus


<400> 617
Gly Thr Thr Tyr Gly Met Cys Thr Glu Lys Phe Thr Phe Ser Lys Asn
 1               5                  10                  15
Pro Ala Asp Thr Gly His Gly Thr Val Val Leu Glu Leu Gln Tyr Thr
            20                  25                  30
Gly Ser Asp Gly Pro Cys Lys Ile Pro Ile Ser Ser Val Ala Ser Leu
        35                  40                  45
Asn Asp Met Thr Pro Val Gly Arg Met Val Thr Ala Asn Pro Tyr Val
    50                  55                  60
Ala Ser Ser Thr Ala Asn Ala Lys Val Leu Val Glu Ile Glu Pro Pro
65                  70                  75                  80
Phe Gly Asp Ser Tyr Ile Val Val Gly Arg Gly Asp Lys Gln Ile Asn
                85                  90                  95
His His Trp His Lys Glu Gly
            100


<210> 618
```

<211> 103
<212> PRT
<213> Artificial Sequence

<220>
<223> Japanese encephalitis virus

<400> 618
Gly Thr Thr Tyr Gly Met Cys Thr Glu Lys Phe Ser Phe Ala Lys Asn
1 5 10 15
Pro Ala Asp Thr Gly His Gly Thr Val Val Ile Glu Leu Ser Tyr Ser
20 25 30
Gly Ser Asp Gly Pro Cys Lys Ile Pro Ile Val Ser Val Ala Ser Leu
35 40 45
Asn Asp Met Thr Pro Val Gly Arg Leu Val Thr Val Asn Pro Phe Val
50 55 60
Ala Thr Ser Ser Ala Asn Ser Lys Val Leu Val Glu Met Glu Pro Pro
65 70 75 80
Phe Gly Ser Asp Tyr Ile Val Val Gly Met Gly Asp Lys Gln Ile Asn
85 90 95
His His Trp His Lys Ala Gly
100

<210> 619
<211> 26
<212> PRT
<213> Artificial Sequence

<220>
<223> Japanese encephalitis virus

<400> 619
Glu Met Glu Pro Pro Phe Gly Asp Ser Tyr Ile Val Val Met Gly Asp
1 5 10 15
Lys Gln Ile Asn His His Trp His Lys Ala
20 25

<210> 620
<211> 95
<212> PRT
<213> Artificial Sequence

<220>
<223> Tick-borne encephalitis virus

<400> 620
Gly Leu Thr Tyr Thr Met Cys Asp Lys Thr Lys Phe Thr Trp Lys Arg
1 5 10 15
Ala Pro Thr Asp Ser Gly His Asp Thr Val Val Met Glu Val Thr Phe
20 25 30
Ser Gly Thr Lys Pro Cys Arg Ile Pro Val Arg Ala Val Ala His Gly
35 40 45
Ser Pro Asp Val Asn Val Ala Met Leu Ile Thr Pro Asn Pro Thr Ile
50 55 60

```
Glu Asn Asn Gly Gly Gly Phe Ile Glu Met Gln Leu Pro Pro Gly Asp
65                  70                  75                  80
Asn Ile Ile Tyr Val Gly Glu Leu Ser His Gln Trp Phe Gln Lys
                85                  90                  95


<210> 621
<211> 94
<212> PRT
<213> Artificial Sequence

<220>
<223> Yellow fever virus


<400> 621
Gly Leu Thr Tyr Thr Met Cys Asp Lys Thr Phe Thr Trp Lys Arg Ala
 1               5                  10                  15
Pro Thr Asp Ser Gly His Asp Thr Val Val Met Glu Val Thr Phe Ser
            20                  25                  30
Gly Thr Lys Pro Cys Arg Ile Pro Val Arg Ala Val Ala His Gly Ser
        35                  40                  45
Pro Asp Val Asn Val Ala Met Leu Ile Thr Pro Asn Pro Thr Ile Glu
    50                  55                  60
Asn Asn Gly Gly Gly Phe Ile Glu Met Gln Leu Pro Pro Gly Asp Asn
65                  70                  75                  80
Ile Ile Tyr Val Gly Glu Leu Ser His Gln Trp Phe Gln Lys
                85                  90


<210> 622
<211> 95
<212> PRT
<213> Artificial Sequence

<220>
<223> Envelope protein flavivirus


<400> 622
Gly Thr Thr Tyr Gly Met Cys Ser Lys Lys Phe Thr Phe Arg Pro Ala
 1               5                  10                  15
Asp Thr Gly His Gly Thr Val Val Leu Glu Leu Gln Tyr Ser Gly Asp
            20                  25                  30
Gly Pro Cys Lys Ile Pro Ile Ser Val Ala Ser Lys Asn Asp Leu Thr
        35                  40                  45
Pro Val Gly Arg Leu Val Thr Val Asn Pro Phe Val Ser Ser Thr Ala
    50                  55                  60
Asn Ala Lys Val Leu Ile Glu Met Glu Pro Pro Phe Gly Asp Ser Tyr
65                  70                  75                  80
Ile Val Val Gly Gly Glu Gln Ile Asn His His Trp His Lys Gly
                85                  90                  95


<210> 623
<211> 27
<212> PRT
<213> Artificial Sequence
```

<220>
<223> Dengue virus

<400> 623
Glu Ala Glu Pro Pro Phe Gly Glu Ser Tyr Ile Val Val Gly Ala Gly
1 5 10 15
Glu Lys Ala Leu Lys Leu Ser Trp Phe Lys Lys
20 25

<210> 624
<211> 27
<212> PRT
<213> Artificial Sequence

<220>
<223> Dengue virus

<400> 624
Glu Thr Glu Pro Pro Phe Gly Glu Ser Tyr Ile Val Val Gly Ala Gly
1 5 10 15
Glu Lys Ala Leu Lys Leu Ser Trp Phe Lys Lys
20 25

<210> 625
<211> 27
<212> PRT
<213> Artificial Sequence

<220>
<223> Dengue virus

<400> 625
Glu Ala Glu Pro Pro Phe Gly Asp Ser Tyr Ile Ile Ile Gly Val Glu
1 5 10 15
Pro Gln Gln Leu Lys Leu Asn Trp Phe Lys Lys
20 25

<210> 626
<211> 27
<212> PRT
<213> Artificial Sequence

<220>
<223> Dengue virus

<400> 626
Glu Ala Glu Pro Pro Phe Gly Glu Ser Asn Ile Val Ile Gly Ile Gly
1 5 10 15
Asp Lys Ala Leu Lys Ile Asn Trp Tyr Lys Lys
20 25

<210> 627

<211> 27
<212> PRT
<213> Artificial Sequence

<220>
<223> Dengue virus

<400> 627
Glu Leu Glu Pro Pro Phe Gly Glu Ser Tyr Ile Val Ile Gly Val Gly
1 5 10 15
Asn Ser Ala Leu Thr Leu His Trp Phe Arg Lys
20 25

<210> 628
<211> 48
<212> DNA
<213> Artificial Sequence

<220>
<223> Linker

<400> 628
aagggcaatt cgaagcttga aggtcaattg gaattcccta ggactagt 48

<210> 629
<211> 16
<212> PRT
<213> Artificial Sequence

<220>
<223> Linker

<400> 629
Lys Gly Asn Ser Lys Leu Glu Gly Gln Leu Glu Phe Pro Arg Thr Ser
1 5 10 15

<210> 630
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Linker

<400> 630
gaattctgca gatatccagc acagtggcgg ccgctc 36

<210> 631
<211> 12
<212> PRT
<213> Artificial Sequence

<220>

<223> Linker

<400> 631
Glu Phe Cys Arg Tyr Pro Ala Gln Trp Arg Pro Leu
1 5 10

<210> 632
<211> 1863
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2.EIII+

<400> 632
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcagaa agcgtatgat 540
gtgaaagata cagcagtaac aacgaaagct tatgccaata atggtactac actggatgta 600
tcgggtcttg atgatgcagc tattaaagcg gctacgggtg gtacgaatgg tacggcttct 660
gtaaccggtg gtgcggttaa atttgacgca gataataaca agtactttgt tactattggt 720
ggctttactg gtgctgatgc cgccaaaaat ggcgattatg aagttaacgt tgctactgac 780
ggtacagtaa cccttgcggc tggcgcaact aaaaccacaa tgcctgctgg tgcgacaact 840
aaaacagaag tacaggagtt aaaagataca ccggcagttg tttcagcaga tgctaaaaat 900
gccttaattg ctggcggcgt tgacgctacc gatgctaatg gcgctgagtt ggtcaaaatg 960
tcttataccg ataaaaatgg taagacaatt gaaggcggtt atgcgcttaa agctggcgat 1020
aagtattacg ccgcagatta cgatgaagcg acaggagcaa ttaaagctaa aactacaagt 1080
tatactgctg ctgacggcac taccaaaaca gcggctaacc aactgggtgg cgtagacggt 1140
aaaaccgaag tcgttactat cgacggtaaa acctacaatg ccagcaaagc cgctggtcat 1200
gatttcaaag cacaaccaga gctggcggaa gcagccgcta aaaccaccga aaacccgctg 1260
cagaaaattg atgccgcgct ggcgcaggtg gatgcgctgc gctctgatct gggtgcggta 1320
caaaaccgtt tcaactctgc tatcaccaac ctgggcaata ccgtaaacaa tctgtctgaa 1380
gcgcgtagcc gtatcgaaga ttccgactac gcgaccgaag tttccaacat gtctcgcgcg 1440
cagattctgc agcaggccgg tacttccgtt ctggcgcagg ctaaccaggt cccgcagaac 1500
gtgctgtctc tgttacgtat ggaaaaattg cagttgaagg gaacaaccta tggcgtctgt 1560
tcaaaggctt tcaagtttct tgggactccc gcagacacag gtcacggcac tgtggtgttg 1620
gaattgcagt acactggcac ggatggacct tgcaaagttc ctatctcgtc agtggcttca 1680
ttgaacgacc taacgccagt gggcagattg gtcactgtca acccttttgt ttcagtggcc 1740
acggccaacg ctaaggtcct gattgaattg gaaccaccct ttggagactc atacatagtg 1800
gtgggcagag gagaacaaca gatcaatcac cattggcaca agtctggaag cagcattggc 1860
aaa 1863

<210> 633
<211> 621
<212> PRT
<213> Artificial Sequence

<220>
<223> STF2.EIII+

<400> 633
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
1 5 10 15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
20 25 30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
35 40 45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
50 55 60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65 70 75 80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
85 90 95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
100 105 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
115 120 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
130 135 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145 150 155 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val Gln
165 170 175
Lys Ala Tyr Asp Val Lys Asp Thr Ala Val Thr Thr Lys Ala Tyr Ala
180 185 190
Asn Asn Gly Thr Thr Leu Asp Val Ser Gly Leu Asp Asp Ala Ala Ile
195 200 205
Lys Ala Ala Thr Gly Gly Thr Asn Gly Thr Ala Ser Val Thr Gly Gly
210 215 220
Ala Val Lys Phe Asp Ala Asp Asn Asn Lys Tyr Phe Val Thr Ile Gly
225 230 235 240
Gly Phe Thr Gly Ala Asp Ala Ala Lys Asn Gly Asp Tyr Glu Val Asn
245 250 255
Val Ala Thr Asp Gly Thr Val Thr Leu Ala Ala Gly Ala Thr Lys Thr
260 265 270
Thr Met Pro Ala Gly Ala Thr Thr Lys Thr Glu Val Gln Glu Leu Lys
275 280 285
Asp Thr Pro Ala Val Val Ser Ala Asp Ala Lys Asn Ala Leu Ile Ala
290 295 300
Gly Gly Val Asp Ala Thr Asp Ala Asn Gly Ala Glu Leu Val Lys Met
305 310 315 320
Ser Tyr Thr Asp Lys Asn Gly Lys Thr Ile Glu Gly Gly Tyr Ala Leu
325 330 335
Lys Ala Gly Asp Lys Tyr Tyr Ala Ala Asp Tyr Asp Glu Ala Thr Gly
340 345 350
Ala Ile Lys Ala Lys Thr Thr Ser Tyr Thr Ala Ala Asp Gly Thr Thr
355 360 365
Lys Thr Ala Ala Asn Gln Leu Gly Gly Val Asp Gly Lys Thr Glu Val
370 375 380
Val Thr Ile Asp Gly Lys Thr Tyr Asn Ala Ser Lys Ala Ala Gly His
385 390 395 400
Asp Phe Lys Ala Gln Pro Glu Leu Ala Glu Ala Ala Ala Lys Thr Thr
405 410 415
Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala
420 425 430
Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile
435 440 445
Thr Asn Leu Gly Asn Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg

```
    450                 455                 460
Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala
465                 470                 475                 480
Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln
                485                 490                 495
Val Pro Gln Asn Val Leu Ser Leu Leu Arg Met Glu Lys Leu Gln Leu
            500                 505                 510
Lys Gly Thr Thr Tyr Gly Val Cys Ser Lys Ala Phe Lys Phe Leu Gly
        515                 520                 525
Thr Pro Ala Asp Thr Gly His Gly Thr Val Val Leu Glu Leu Gln Tyr
    530                 535                 540
Thr Gly Thr Asp Gly Pro Cys Lys Val Pro Ile Ser Ser Val Ala Ser
545                 550                 555                 560
Leu Asn Asp Leu Thr Pro Val Gly Arg Leu Val Thr Val Asn Pro Phe
                565                 570                 575
Val Ser Val Ala Thr Ala Asn Ala Lys Val Leu Ile Glu Leu Glu Pro
            580                 585                 590
Pro Phe Gly Asp Ser Tyr Ile Val Val Gly Arg Gly Glu Gln Gln Ile
        595                 600                 605
Asn His His Trp His Lys Ser Gly Ser Ser Ile Gly Lys
    610                 615                 620
```

<210> 634
<211> 1176
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2delta.EIII+

<400> 634

```
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcatgg agcgccggtg 540
gatcctgcta gcccatggac cgaaaacccg ctgcagaaaa ttgatgccgc gctggcgcag 600
gtggatgcgc tgcgctctga tctgggtgcg gtacaaaacc gtttcaactc tgctatcacc 660
aacctgggca ataccgtaaa caatctgtct gaagcgcgta gccgtatcga agattccgac 720
tacgcgaccg aagtttccaa catgtctcgc gcgcagattt tgcagcaggc cggtacttcc 780
gttctggcgc aggctaacca ggtcccgcag aacgtgctgt ctctgttacg tatggaaaaa 840
ttgcagttga agggaacaac ctatggcgtc tgttcaaagg ctttcaagtt tcttgggact 900
cccgcagaca caggtcacgg cactgtggtg ttggaattgc agtacactgg cacggatgga 960
ccttgcaaag ttcctatctc gtcagtggct tcattgaacg acctaacgcc agtgggcaga 1020
ttggtcactg tcaacccttt tgtttcagtg gccacggcca acgctaaggt cctgattgaa 1080
ttggaaccac cctttggaga ctcatacata gtggtgggca gaggagaaca acagatcaat 1140
caccattggc acaagtctgg aagcagcatt ggcaaa                          1176
```

<210> 635
<211> 392
<212> PRT
<213> Artificial Sequence

<220>
<223> STF2delta.EIII+

<400> 635

```
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
 1               5                  10                  15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
            20                  25                  30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
        35                  40                  45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
    50                  55                  60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65                  70                  75                  80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
                85                  90                  95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
            100                 105                 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
        115                 120                 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
    130                 135                 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145                 150                 155                 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val His
                165                 170                 175
Gly Ala Pro Val Asp Pro Ala Ser Pro Trp Thr Glu Asn Pro Leu Gln
            180                 185                 190
Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala Leu Arg Ser Asp Leu
        195                 200                 205
Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile Thr Asn Leu Gly Asn
    210                 215                 220
Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg Ile Glu Asp Ser Asp
225                 230                 235                 240
Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala Gln Ile Leu Gln Gln
                245                 250                 255
Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln Val Pro Gln Asn Val
            260                 265                 270
Leu Ser Leu Leu Arg Met Glu Lys Leu Gln Leu Lys Gly Thr Thr Tyr
        275                 280                 285
Gly Val Cys Ser Lys Ala Phe Lys Phe Leu Gly Thr Pro Ala Asp Thr
    290                 295                 300
Gly His Gly Thr Val Val Leu Glu Leu Gln Tyr Thr Gly Thr Asp Gly
305                 310                 315                 320
Pro Cys Lys Val Pro Ile Ser Ser Val Ala Ser Leu Asn Asp Leu Thr
                325                 330                 335
Pro Val Gly Arg Leu Val Thr Val Asn Pro Phe Val Ser Val Ala Thr
            340                 345                 350
Ala Asn Ala Lys Val Leu Ile Glu Leu Glu Pro Pro Phe Gly Asp Ser
        355                 360                 365
Tyr Ile Val Val Gly Arg Gly Glu Gln Gln Ile Asn His His Trp His
    370                 375                 380
Lys Ser Gly Ser Ser Ile Gly Lys
385                 390
```

<210> 636
<211> 1224

<212> DNA
<213> Artificial Sequence

<220>
<223> STF2delta.EIII+

<400> 636
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcatgg agcgccggtg 540
gatcctgcta gcccatggac cgaaaacccg ctgcagaaaa ttgatgccgc gctggcgcag 600
gtggatgcgc tgcgctctga tctgggtgcg gtacaaaacc gtttcaactc tgctatcacc 660
aacctgggca ataccgtaaa caatctgtct gaagcgcgta gccgtatcga agattccgac 720
tacgcgaccg aagtttccaa catgtctcgc gcgcagattt tgcagcaggc cggtacttcc 780
gttctggcgc aggctaacca ggtcccgcag aacgtgctgt ctctgttacg tgaattctgc 840
agatatccag cacagtggcg gccgctcatg gaaaaattgc agttgaaggg aacaacctat 900
ggcgtctgtt caaaggcttt caagtttctt gggactcccg cagacacagg tcacggcact 960
gtggtgttgg aattgcagta cactggcacg gatggacctt gcaaagttcc tatctcgtca 1020
gtggcttcat tgaacgacct aacgccagtg ggcagattgg tcactgtcaa ccctttttgtt 1080
tcagtggcca cggccaacgc taaggtcctg attgaattgg aaccaccctt tggagactca 1140
tacatagtgg tgggcagagg agaacaacag atcaatcacc attggcacaa gtctggaagc 1200
agcattggca aacccttaat aagc 1224

<210> 637
<211> 408
<212> PRT
<213> Artificial Sequence

<220>
<223> STF2delta.EIII+

<400> 637
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
1 5 10 15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
20 25 30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
35 40 45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
50 55 60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65 70 75 80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
85 90 95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
100 105 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
115 120 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
130 135 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu

```
145                 150                 155                 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val His
                165                 170                 175
Gly Ala Pro Val Asp Pro Ala Ser Pro Trp Thr Glu Asn Pro Leu Gln
            180                 185                 190
Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala Leu Arg Ser Asp Leu
        195                 200                 205
Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile Thr Asn Leu Gly Asn
    210                 215                 220
Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg Ile Glu Asp Ser Asp
225                 230                 235                 240
Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala Gln Ile Leu Gln Gln
                245                 250                 255
Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln Val Pro Gln Asn Val
            260                 265                 270
Leu Ser Leu Leu Arg Glu Phe Cys Arg Tyr Pro Ala Gln Trp Arg Pro
        275                 280                 285
Leu Met Glu Lys Leu Gln Leu Lys Gly Thr Thr Tyr Gly Val Cys Ser
    290                 295                 300
Lys Ala Phe Lys Phe Leu Gly Thr Pro Ala Asp Thr Gly His Gly Thr
305                 310                 315                 320
Val Val Leu Glu Leu Gln Tyr Thr Gly Thr Asp Gly Pro Cys Lys Val
                325                 330                 335
Pro Ile Ser Ser Val Ala Ser Leu Asn Asp Leu Thr Pro Val Gly Arg
            340                 345                 350
Leu Val Thr Val Asn Pro Phe Val Ser Val Ala Thr Ala Asn Ala Lys
        355                 360                 365
Val Leu Ile Glu Leu Glu Pro Pro Phe Gly Asp Ser Tyr Ile Val Val
    370                 375                 380
Gly Arg Gly Glu Gln Gln Ile Asn His His Trp His Lys Ser Gly Ser
385                 390                 395                 400
Ser Ile Gly Lys Pro Leu Ile Ser
                405
```

<210> 638
<211> 1899
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2.EIII+

<400> 638

```
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcagaa agcgtatgat 540
gtgaaagata cagcagtaac aacgaaagct tatgccaata atggtactac actggatgta 600
tcgggtcttg atgatgcagc tattaaagcg gctacgggtg gtacgaatgg tacggcttct 660
gtaaccggtg gtgcggttaa atttgacgca gataataaca agtactttgt tactattggt 720
ggctttactg gtgctgatgc cgccaaaaat ggcgattatg aagttaacgt tgctactgac 780
ggtacagtaa cccttgcggc tggcgcaact aaaaccacaa tgcctgctgg tgcgacaact 840
```

```
aaaacagaag tacaggagtt aaaagataca ccggcagttg tttcagcaga tgctaaaaat 900
gccttaattg ctggcggcgt tgacgctacc gatgctaatg gcgctgagtt ggtcaaaatg 960
tcttataccg ataaaaatgg taagacaatt gaaggcggtt atgcgcttaa agctggcgat 1020
aagtattacg ccgcagatta cgatgaagcg acaggagcaa ttaaagctaa aactacaagt 1080
tatactgctg ctgacggcac taccaaaaca gcggctaacc aactgggtgg cgtagacggt 1140
aaaaccgaag tcgttactat cgacggtaaa acctacaatg ccagcaaagc cgctggtcat 1200
gatttcaaag cacaaccaga gctggcggaa gcagccgcta aaaccaccga aaacccgctg 1260
cagaaaattg atgccgcgct ggcgcaggtg gatgcgctgc gctctgatct gggtgcggta 1320
caaaaccgtt tcaactctgc tatcaccaac ctgggcaata ccgtaaacaa tctgtctgaa 1380
gcgcgtagcc gtatcgaaga ttccgactac gcgaccgaag tttccaacat gtctcgcgcg 1440
cagattctgc agcaggccgg tacttccgtt ctggcgcagg ctaaccaggt cccgcagaac 1500
gtgctgtctc tgttacgtga attctgcaga tatccagcac agtggcggcc gctcatggaa 1560
aaattgcagt tgaagggaac aacctatggc gtctgttcaa aggctttcaa gtttcttggg 1620
actcccgcag acacaggtca cggcactgtg gtgttggaat tgcagtacac tggcacggat 1680
ggaccttgca aagttcctat ctcgtcagtg gcttcattga acgacctaac gccagtgggc 1740
agattggtca ctgtcaaccc ttttgtttca gtggccacgg ccaacgctaa ggtcctgatt 1800
gaattggaac caccctttgg agactcatac atagtggtgg gcagaggaga acaacagatc 1860
aatcaccatt ggcacaagtc tggaagcagc attggcaaa                        1899

<210> 639
<211> 633
<212> PRT
<213> Artificial Sequence

<220>
<223> STF2.EIII+


<400> 639
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
 1               5                  10                  15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
            20                  25                  30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
        35                  40                  45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
    50                  55                  60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65                  70                  75                  80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
                85                  90                  95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
            100                 105                 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
        115                 120                 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
    130                 135                 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145                 150                 155                 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val Gln
                165                 170                 175
Lys Ala Tyr Asp Val Lys Asp Thr Ala Val Thr Thr Lys Ala Tyr Ala
            180                 185                 190
Asn Asn Gly Thr Thr Leu Asp Val Ser Gly Leu Asp Asp Ala Ala Ile
        195                 200                 205
Lys Ala Ala Thr Gly Gly Thr Asn Gly Thr Ala Ser Val Thr Gly Gly
    210                 215                 220
Ala Val Lys Phe Asp Ala Asp Asn Asn Lys Tyr Phe Val Thr Ile Gly
225                 230                 235                 240
```

```
Gly Phe Thr Gly Ala Asp Ala Ala Lys Asn Gly Asp Tyr Glu Val Asn
                245                 250                 255
Val Ala Thr Asp Gly Thr Val Thr Leu Ala Ala Gly Ala Thr Lys Thr
            260                 265                 270
Thr Met Pro Ala Gly Ala Thr Thr Lys Thr Glu Val Gln Glu Leu Lys
        275                 280                 285
Asp Thr Pro Ala Val Val Ser Ala Asp Ala Lys Asn Ala Leu Ile Ala
    290                 295                 300
Gly Gly Val Asp Ala Thr Asp Ala Asn Gly Ala Glu Leu Val Lys Met
305                 310                 315                 320
Ser Tyr Thr Asp Lys Asn Gly Lys Thr Ile Glu Gly Gly Tyr Ala Leu
                325                 330                 335
Lys Ala Gly Asp Lys Tyr Tyr Ala Ala Asp Tyr Asp Glu Ala Thr Gly
            340                 345                 350
Ala Ile Lys Ala Lys Thr Thr Ser Tyr Thr Ala Ala Asp Gly Thr Thr
        355                 360                 365
Lys Thr Ala Ala Asn Gln Leu Gly Gly Val Asp Gly Lys Thr Glu Val
    370                 375                 380
Val Thr Ile Asp Gly Lys Thr Tyr Asn Ala Ser Lys Ala Ala Gly His
385                 390                 395                 400
Asp Phe Lys Ala Gln Pro Glu Leu Ala Glu Ala Ala Ala Lys Thr Thr
                405                 410                 415
Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala
            420                 425                 430
Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile
        435                 440                 445
Thr Asn Leu Gly Asn Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg
    450                 455                 460
Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala
465                 470                 475                 480
Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln
                485                 490                 495
Val Pro Gln Asn Val Leu Ser Leu Leu Arg Glu Phe Cys Arg Tyr Pro
            500                 505                 510
Ala Gln Trp Arg Pro Leu Met Glu Lys Leu Gln Leu Lys Gly Thr Thr
        515                 520                 525
Tyr Gly Val Cys Ser Lys Ala Phe Lys Phe Leu Gly Thr Pro Ala Asp
    530                 535                 540
Thr Gly His Gly Thr Val Val Leu Glu Leu Gln Tyr Thr Gly Thr Asp
545                 550                 555                 560
Gly Pro Cys Lys Val Pro Ile Ser Ser Val Ala Ser Leu Asn Asp Leu
                565                 570                 575
Thr Pro Val Gly Arg Leu Val Thr Val Asn Pro Phe Val Ser Val Ala
            580                 585                 590
Thr Ala Asn Ala Lys Val Leu Ile Glu Leu Glu Pro Pro Phe Gly Asp
        595                 600                 605
Ser Tyr Ile Val Val Gly Arg Gly Glu Gln Gln Ile Asn His His Trp
    610                 615                 620
His Lys Ser Gly Ser Ser Ile Gly Lys
625                 630
```

<210> 640
<211> 1143
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2delta.EIII+

<400> 640
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgaccga aaacccgctg 540
cagaaaattg atgccgcgct ggcgcaggtg gatgcgctgc gctctgatct gggtgcggta 600
caaaaccgtt tcaactctgc tatcaccaac ctgggcaata ccgtaaacaa tctgtctgaa 660
gcgcgtagcc gtatcgaaga ttccgactac gcgaccgaag tttccaacat gtctcgcgcg 720
cagattctgc agcaggccgg tacttccgtt ctggcgcagg ctaaccaggt cccgcagaac 780
gtgctgtctc tgttacgtat ggaaaaattg cagttgaagg gaacaaccta tggcgtctgt 840
tcaaaggctt tcaagtttct tgggactccc gcagacacag gtcacggcac tgtggtgttg 900
gaattgcagt acactggcac ggatggacct tgcaaagttc ctatctcgtc agtggcttca 960
ttgaacgacc taacgccagt gggcagattg gtcactgtca acccttttgt ttcagtggcc 1020
acggccaacg ctaaggtcct gattgaattg gaaccaccct ttggagactc atacatagtg 1080
gtgggcagag gagaacaaca gatcaatcac cattggcaca agtctggaag cagcattggc 1140
aaa 1143

<210> 641
<211> 381
<212> PRT
<213> Artificial Sequence

<220>
<223> STF2delta.EIII+

<400> 641
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
1 5 10 15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
20 25 30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
35 40 45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
50 55 60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65 70 75 80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
85 90 95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
100 105 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
115 120 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
130 135 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145 150 155 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val Thr
165 170 175
Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala
180 185 190
Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile

195 200 205
Thr Asn Leu Gly Asn Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg
210 215 220
Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala
225 230 235 240
Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln
245 250 255
Val Pro Gln Asn Val Leu Ser Leu Leu Arg Met Glu Lys Leu Gln Leu
260 265 270
Lys Gly Thr Thr Tyr Gly Val Cys Ser Lys Ala Phe Lys Phe Leu Gly
275 280 285
Thr Pro Ala Asp Thr Gly His Gly Thr Val Val Leu Glu Leu Gln Tyr
290 295 300
Thr Gly Thr Asp Gly Pro Cys Lys Val Pro Ile Ser Ser Val Ala Ser
305 310 315 320
Leu Asn Asp Leu Thr Pro Val Gly Arg Leu Val Thr Val Asn Pro Phe
325 330 335
Val Ser Val Ala Thr Ala Asn Ala Lys Val Leu Ile Glu Leu Glu Pro
340 345 350
Pro Phe Gly Asp Ser Tyr Ile Val Val Gly Arg Gly Glu Gln Gln Ile
355 360 365
Asn His His Trp His Lys Ser Gly Ser Ser Ile Gly Lys
370 375 380

<210> 642
<211> 406
<212> PRT
<213> Artificial Sequence

<220>
<223> West Nile virus envelope

<400> 642
Phe Asn Cys Leu Gly Met Ser Asn Arg Asp Phe Leu Glu Gly Val Ser
1 5 10 15
Gly Ala Thr Trp Val Asp Leu Val Leu Glu Gly Asp Ser Cys Val Thr
20 25 30
Ile Met Ser Lys Asp Lys Pro Thr Ile Asp Val Lys Met Met Asn Met
35 40 45
Glu Ala Ala Asn Leu Ala Glu Val Arg Ser Tyr Cys Tyr Leu Ala Thr
50 55 60
Val Ser Asp Leu Ser Thr Lys Ala Ala Cys Pro Thr Met Gly Glu Ala
65 70 75 80
His Asn Asp Lys Arg Ala Asp Pro Ala Phe Val Cys Arg Gln Gly Val
85 90 95
Val Asp Arg Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly Ser
100 105 110
Ile Asp Thr Cys Ala Lys Phe Ala Cys Ser Thr Lys Ala Ile Gly Arg
115 120 125
Thr Ile Leu Lys Glu Asn Ile Lys Tyr Glu Val Ala Ile Phe Val His
130 135 140
Gly Pro Thr Thr Val Glu Ser His Gly Asn Tyr Ser Thr Gln Val Gly
145 150 155 160
Ala Thr Gln Ala Gly Arg Phe Ser Ile Thr Pro Ala Ala Pro Ser Tyr
165 170 175
Thr Leu Lys Leu Gly Glu Tyr Gly Glu Val Thr Val Asp Cys Glu Pro
180 185 190

```
Arg Ser Gly Ile Asp Thr Asn Ala Tyr Tyr Val Met Thr Val Gly Thr
        195                 200                 205
Lys Thr Phe Leu Val His Arg Glu Trp Phe Met Asp Leu Asn Leu Pro
    210                 215                 220
Trp Ser Ser Ala Gly Ser Thr Val Trp Arg Asn Arg Glu Thr Leu Met
225                 230                 235                 240
Glu Phe Glu Glu Pro His Ala Thr Lys Gln Ser Val Ile Ala Leu Gly
                245                 250                 255
Ser Gln Glu Gly Ala Leu His Gln Ala Leu Ala Gly Ala Ile Pro Val
            260                 265                 270
Glu Phe Ser Ser Asn Thr Val Lys Leu Thr Ser Gly His Leu Lys Cys
        275                 280                 285
Arg Val Lys Met Glu Lys Leu Gln Leu Lys Gly Thr Thr Tyr Gly Val
    290                 295                 300
Cys Ser Lys Ala Phe Lys Phe Leu Gly Thr Pro Ala Asp Thr Gly His
305                 310                 315                 320
Gly Thr Val Val Leu Glu Leu Gln Tyr Thr Gly Thr Asp Gly Pro Cys
                325                 330                 335
Lys Val Pro Ile Ser Ser Val Ala Ser Leu Asn Asp Leu Thr Pro Val
            340                 345                 350
Gly Arg Leu Val Thr Val Asn Pro Phe Val Ser Val Ala Thr Ala Asn
        355                 360                 365
Ala Lys Val Leu Ile Glu Leu Glu Pro Pro Phe Gly Asp Ser Tyr Ile
    370                 375                 380
Val Val Gly Arg Gly Glu Gln Gln Ile Asn His His Trp His Lys Ser
385                 390                 395                 400
Gly Ser Ser Ile Gly Lys
                405
```

<210> 643
<211> 99
<212> PRT
<213> Artificial Sequence

<220>
<223> flavivirus

```
<400> 643
Gly Met Ser Tyr Ser Met Cys Thr Gly Lys Phe Lys Val Val Lys Glu
 1                   5                  10                  15
Ile Ala Glu Thr Gln His Gly Thr Ile Val Ile Arg Val Gln Tyr Glu
            20                  25                  30
Gly Asp Gly Ser Pro Cys Lys Ile Pro Phe Glu Ile Met Asp Leu Glu
        35                  40                  45
Lys Arg His Val Leu Gly Arg Leu Ile Thr Val Asn Pro Ile Val Thr
    50                  55                  60
Glu Lys Asp Ser Pro Val Asn Ile Glu Ala Glu Pro Pro Phe Gly Asp
65                  70                  75                  80
Ser Tyr Ile Ile Ile Gly Val Glu Pro Gly Gln Leu Lys Leu Asn Trp
                85                  90                  95
Phe Lys Lys
```

<210> 644
<211> 38
<212> DNA

<213> Artificial Sequence

<220>
<223> primer

<400> 644
ctcgggagat ctgcacaagt aatcaacact aacagtct 38

<210> 645
<211> 41
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 645
ccatgggcta gcaggatcca ccggcgctcc ctgcacgttc a 41

<210> 646
<211> 42
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 646
ggagcgccgg tggatcctgc tagcccatgg accgaaaacc cg 42

<210> 647
<211> 84
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 647
tctgcagaat tcacgtaaca gagacagcac gttctgcggg acgtcccgca gaacgtgctg 60
tctctgttac gtgaattctg caga 84

<210> 648
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 648
tccggcgtag aggatcgaga 20

<210> 649

<211> 54
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 649
caattgacct tcaagcttcg aattgcccct acgtaacaga gacagcacgt tctg 54

<210> 650
<211> 57
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 650
aagcttgaag gtcaattgga attccctagg actagtatgg aaaaattgca gttgaag 57

<210> 651
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 651
gcttaatgcg ccgctacagg 20

<210> 652
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 652
gcggccgctc atggaaaaat tgcagttgaa gggaacaacc 40

<210> 653
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 653
ccgcggtttg ccaatgctgc ttccagactt gt 32

<210> 654

<211> 49
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 654
ccggcatgcc atatggcaca agtaatcaac actaacagtc tgtcgctgc 49

<210> 655
<211> 49
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 655
gcatgctcag cttattaagg gtttgccaat gctgcttccc agacttgtg 49

<210> 656
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 656
tacgtgaatt cagcagatat ccagcac 27

<210> 657
<211> 1911
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2.EIII+

<400> 657
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcagaa agcgtatgat 540
gtgaaagata cagcagtaac aacgaaagct tatgccaata atggtactac actggatgta 600
tcgggtcttg atgatgcagc tattaaagcg gctacgggtg gtacgaatgg tacggcttct 660
gtaaccggtg gtgcggttaa atttgacgca gataataaca agtactttgt tactattggt 720
ggctttactg gtgctgatgc cgccaaaaat ggcgattatg aagttaacgt tgctactgac 780
ggtacagtaa cccttgcggc tggcgcaact aaaaccacaa tgcctgctgg tgcgacaact 840

aaaacagaag tacaggagtt aaaagataca ccggcagttg tttcagcaga tgctaaaaat 900
gccttaattg ctggcggcgt tgacgctacc gatgctaatg gcgctgagtt ggtcaaaatg 960
tcttataccg ataaaaatgg taagacaatt gaaggcggtt atgcgcttaa agctggcgat 1020
aagtattacg ccgcagatta cgatgaagcg acaggagcaa ttaaagctaa aactacaagt 1080
tatactgctg ctgacggcac taccaaaaca gcggctaacc aactgggtgg cgtagacggt 1140
aaaaccgaag tcgttactat cgacggtaaa acctacaatg ccagcaaagc cgctggtcat 1200
gatttcaaag cacaaccaga gctggcggaa gcagccgcta aaaccaccga aaacccgctg 1260
cagaaaattg atgccgcgct ggcgcaggtg gatgcgctgc gctctgatct gggtgcggta 1320
caaaaccgtt tcaactctgc tatcaccaac ctgggcaata ccgtaaacaa tctgtctgaa 1380
gcgcgtagcc gtatcgaaga ttccgactac gcgaccgaag tttccaacat gtctcgcgcg 1440
cagattctgc agcaggccgg tacttccgtt ctggcgcagg ctaaccaggt cccgcagaac 1500
gtgctgtctc tgttacgtaa gggcaattcg aagcttgaag gtcaattgga attccctagg 1560
actagtatgg aaaaattgca gttgaaggga acaacctatg gcgtctgttc aaaggctttc 1620
aagtttcttg ggactcccgc agacacaggt cacggcactg tggtgttgga attgcagtac 1680
actggcacgg atggaccttg caaagttcct atctcgtcag tggcttcatt gaacgaccta 1740
acgccagtgg gcagattggt cactgtcaac ccttttgttt cagtggccac ggccaacgct 1800
aaggtcctga ttgaattgga accacccttt ggagactcat acatagtggt gggcagagga 1860
gaacaacaga tcaatcacca ttggcacaag tctggaagca gcattggcaa a 1911

<210> 658
<211> 637
<212> PRT
<213> Artificial Sequence

<220>
<223> STF2.EIII+

<400> 658

Met	Ala	Gln	Val	Ile	Asn	Thr	Asn	Ser	Leu	Ser	Leu	Leu	Thr	Gln	Asn
1				5					10					15	
Asn	Leu	Asn	Lys	Ser	Gln	Ser	Ala	Leu	Gly	Thr	Ala	Ile	Glu	Arg	Leu
			20					25					30		
Ser	Ser	Gly	Leu	Arg	Ile	Asn	Ser	Ala	Lys	Asp	Asp	Ala	Ala	Gly	Gln
		35					40					45			
Ala	Ile	Ala	Asn	Arg	Phe	Thr	Ala	Asn	Ile	Lys	Gly	Leu	Thr	Gln	Ala
	50					55					60				
Ser	Arg	Asn	Ala	Asn	Asp	Gly	Ile	Ser	Ile	Ala	Gln	Thr	Thr	Glu	Gly
65					70					75					80
Ala	Leu	Asn	Glu	Ile	Asn	Asn	Asn	Leu	Gln	Arg	Val	Arg	Glu	Leu	Ala
				85					90					95	
Val	Gln	Ser	Ala	Asn	Ser	Thr	Asn	Ser	Gln	Ser	Asp	Leu	Asp	Ser	Ile
			100					105					110		
Gln	Ala	Glu	Ile	Thr	Gln	Arg	Leu	Asn	Glu	Ile	Asp	Arg	Val	Ser	Gly
		115					120					125			
Gln	Thr	Gln	Phe	Asn	Gly	Val	Lys	Val	Leu	Ala	Gln	Asp	Asn	Thr	Leu
	130					135					140				
Thr	Ile	Gln	Val	Gly	Ala	Asn	Asp	Gly	Glu	Thr	Ile	Asp	Ile	Asp	Leu
145					150					155					160
Lys	Gln	Ile	Asn	Ser	Gln	Thr	Leu	Gly	Leu	Asp	Ser	Leu	Asn	Val	Gln
				165					170					175	
Lys	Ala	Tyr	Asp	Val	Lys	Asp	Thr	Ala	Val	Thr	Thr	Lys	Ala	Tyr	Ala
			180					185					190		
Asn	Asn	Gly	Thr	Thr	Leu	Asp	Val	Ser	Gly	Leu	Asp	Asp	Ala	Ala	Ile
		195					200					205			
Lys	Ala	Ala	Thr	Gly	Gly	Thr	Asn	Gly	Thr	Ala	Ser	Val	Thr	Gly	Gly
	210					215					220				
Ala	Val	Lys	Phe	Asp	Ala	Asp	Asn	Asn	Lys	Tyr	Phe	Val	Thr	Ile	Gly
225					230					235					240

```
Gly Phe Thr Gly Ala Asp Ala Ala Lys Asn Gly Asp Tyr Glu Val Asn
                245                 250                 255
Val Ala Thr Asp Gly Thr Val Thr Leu Ala Ala Gly Ala Thr Lys Thr
            260                 265                 270
Thr Met Pro Ala Gly Ala Thr Thr Lys Thr Glu Val Gln Glu Leu Lys
        275                 280                 285
Asp Thr Pro Ala Val Val Ser Ala Asp Ala Lys Asn Ala Leu Ile Ala
    290                 295                 300
Gly Gly Val Asp Ala Thr Asp Ala Asn Gly Ala Glu Leu Val Lys Met
305                 310                 315                 320
Ser Tyr Thr Asp Lys Asn Gly Lys Thr Ile Glu Gly Gly Tyr Ala Leu
                325                 330                 335
Lys Ala Gly Asp Lys Tyr Tyr Ala Ala Asp Tyr Asp Glu Ala Thr Gly
            340                 345                 350
Ala Ile Lys Ala Lys Thr Thr Ser Tyr Thr Ala Ala Asp Gly Thr Thr
        355                 360                 365
Lys Thr Ala Ala Asn Gln Leu Gly Gly Val Asp Gly Lys Thr Glu Val
    370                 375                 380
Val Thr Ile Asp Gly Lys Thr Tyr Asn Ala Ser Lys Ala Ala Gly His
385                 390                 395                 400
Asp Phe Lys Ala Gln Pro Glu Leu Ala Glu Ala Ala Ala Lys Thr Thr
                405                 410                 415
Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala
            420                 425                 430
Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile
        435                 440                 445
Thr Asn Leu Gly Asn Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg
    450                 455                 460
Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala
465                 470                 475                 480
Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln
                485                 490                 495
Val Pro Gln Asn Val Leu Ser Leu Leu Arg Lys Gly Asn Ser Lys Leu
            500                 505                 510
Glu Gly Gln Leu Glu Phe Pro Arg Thr Ser Met Glu Lys Leu Gln Leu
        515                 520                 525
Lys Gly Thr Thr Tyr Gly Val Cys Ser Lys Ala Phe Lys Phe Leu Gly
    530                 535                 540
Thr Pro Ala Asp Thr Gly His Gly Thr Val Val Leu Glu Leu Gln Tyr
545                 550                 555                 560
Thr Gly Thr Asp Gly Pro Cys Lys Val Pro Ile Ser Ser Val Ala Ser
                565                 570                 575
Leu Asn Asp Leu Thr Pro Val Gly Arg Leu Val Thr Val Asn Pro Phe
            580                 585                 590
Val Ser Val Ala Thr Ala Asn Ala Lys Val Leu Ile Glu Leu Glu Pro
        595                 600                 605
Pro Phe Gly Asp Ser Tyr Ile Val Val Gly Arg Gly Glu Gln Gln Ile
    610                 615                 620
Asn His His Trp His Lys Ser Gly Ser Ser Ile Gly Lys
625                 630                 635


<210> 659
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> Linker
```

<400> 659
Gly Ala Pro Val Asp Pro Ala Ser Pro Trp
1 5 10

<210> 660
<211> 1218
<212> DNA
<213> Artificial Sequence

<220>
<223> West Nile virus envelope protein

<400> 660
ttcaactgcc ttggaatgag caacagagac ttcttggaag gagtgtctgg agcaacatgg 60
gtggatttgg ttctcgaagg cgacagctgc gtgactatca tgtctaagga caagcctacc 120
atcgatgtga agatgatgaa tatggaggcg gccaacctgg cagaggtccg cagttattgc 180
tatttggcta ccgtcagcga tctctccacc aaagctgcgt gcccgaccat gggagaagct 240
cacaatgaca aacgtgctga cccagctttt gtgtgcagac aaggagtggt ggacaggggc 300
tggggcaacg gctgcggact atttggcaaa ggaagcattg acacatgcgc caaatttgcc 360
tgctctacca aggcaatagg aagaaccatc ttgaaagaga atatcaagta cgaagtggcc 420
atttttgtcc atggaccaac tactgtggag tcgcacggaa actactccac acaggttgga 480
gccactcagg cagggagatt cagcatcact cctgcagcgc cttcatacac actaaagctt 540
ggagaatatg gagaggtgac agtggactgt gaaccacggt cagggattga caccaatgca 600
tactacgtga tgactgttgg aacaaagacg ttcttggtcc atcgtgagtg gttcatggac 660
ctcaacctcc cttggagcag tgctggaagt actgtgtgga ggaacagaga gacgttaatg 720
gagtttgagg aaccacacgc cacgaagcag tctgtgatag cattgggctc acaagaggga 780
gctctgcatc aagctttggc tggagccatt cctgtggaat tttcaagcaa cactgtcaag 840
ttgacgtcgg gtcatttgaa gtgtagagtg aagatggaaa aattgcagtt gaagggaaca 900
acctatggcg tctgttcaaa ggctttcaag tttcttggga ctcccgcaga cacaggtcac 960
ggcactgtgg tgttggaatt gcagtacact ggcacggatg gaccttgcaa agttcctatc 1020
tcgtcagtgg cttcattgaa cgacctaacg ccagtgggca gattggtcac tgtcaaccct 1080
tttgtttcag tggccacggc caacgctaag gtcctgattg aattggaacc acccttgga 1140
gactcataca tagtggtggg cagaggagaa caacagatca atcaccattg gcacaagtct 1200
ggaagcagca ttggcaaa 1218

<210> 661
<211> 506
<212> PRT
<213> S. muenchen

<400> 661
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
1 5 10 15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
20 25 30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
35 40 45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
50 55 60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65 70 75 80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
85 90 95
Val Gln Ser Ala Asn Gly Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
100 105 110

```
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
        115                 120                 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
    130                 135                 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145                 150                 155                 160
Lys Glu Ile Ser Ser Lys Thr Leu Gly Leu Asp Lys Leu Asn Val Gln
                165                 170                 175
Asp Ala Tyr Thr Pro Lys Glu Thr Ala Val Thr Val Asp Lys Thr Thr
            180                 185                 190
Tyr Lys Asn Gly Thr Asp Thr Ile Thr Ala Gln Ser Asn Thr Asp Ile
        195                 200                 205
Gln Thr Ala Ile Gly Gly Gly Ala Thr Gly Val Thr Gly Ala Asp Ile
    210                 215                 220
Lys Phe Lys Asp Gly Gln Tyr Tyr Leu Asp Val Lys Gly Gly Ala Ser
225                 230                 235                 240
Ala Gly Val Tyr Lys Ala Thr Tyr Asp Glu Thr Thr Lys Lys Val Asn
                245                 250                 255
Ile Asp Thr Thr Asp Lys Thr Pro Leu Ala Thr Ala Glu Ala Thr Ala
            260                 265                 270
Ile Arg Gly Thr Ala Thr Ile Thr His Asn Gln Ile Ala Glu Val Thr
        275                 280                 285
Lys Glu Gly Val Asp Thr Thr Thr Val Ala Ala Gln Leu Ala Ala Ala
    290                 295                 300
Gly Val Thr Gly Ala Asp Lys Asp Asn Thr Ser Leu Val Lys Leu Ser
305                 310                 315                 320
Phe Glu Asp Lys Asn Gly Lys Val Ile Asp Gly Gly Tyr Ala Val Lys
                325                 330                 335
Met Gly Asp Asp Phe Tyr Ala Ala Thr Tyr Asp Glu Lys Thr Gly Thr
            340                 345                 350
Ile Thr Ala Lys Thr Thr Thr Tyr Thr Asp Gly Ala Gly Val Ala Gln
        355                 360                 365
Thr Gly Ala Val Lys Phe Gly Gly Ala Asn Gly Lys Ser Glu Val Val
    370                 375                 380
Thr Ala Thr Asp Gly Lys Thr Tyr Leu Ala Ser Asp Leu Asp Lys His
385                 390                 395                 400
Asn Phe Arg Thr Gly Gly Glu Leu Lys Glu Val Asn Thr Asp Lys Thr
                405                 410                 415
Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Thr
            420                 425                 430
Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile
        435                 440                 445
Thr Asn Leu Gly Asn Thr Val Asn Asn Leu Ser Ser Ala Arg Ser Arg
    450                 455                 460
Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala
465                 470                 475                 480
Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln
                485                 490                 495
Val Pro Gln Asn Val Leu Ser Leu Leu Arg
            500                 505
```

<210> 662
<211> 1521
<212> DNA
<213> S. muenchen

<400> 662
atggcacaag tcattaatac aaacagcctg tcgctgttga cccagaataa cctgaacaaa 60

tcccagtccg ctctgggcac cgctatcgag cgtctgtctt ccggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacggtacta actcccagtc tgaccttgac tctatccagg ctgaaatcac ccagcgtctg 360
aacgaaatcg accgtgtatc cggtcagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggtgcc aacgacggtg aaactattga tattgattta 480
aaagaaatta gctctaaaac actgggactt gataagctta atgtccagga tgcctacacc 540
ccgaaagaaa ctgctgtaac cgttgataaa actacctata aaaatggtac agatactatt 600
acagcccaga gcaatactga tatccaaact gcaattggcg gtggtgcaac gggggttact 660
ggggctgata tcaaatttaa agatggtcaa tactatttag atgttaaagg cggtgcttct 720
gctggtgttt ataaagccac ttatgatgaa actacaaaga aagttaatat tgatacgact 780
gataaaactc cgttagcaac tgcggaagct acagctattc ggggaacggc cactataacc 840
cacaaccaaa ttgctgaagt aacaaaagag ggtgttgata cgaccacagt tgcggctcaa 900
cttgctgctg caggggttac tggtgccgat aaggacaata ctagccttgt aaaactatcg 960
tttgaggata aaaacggtaa ggttattgat ggtggctatg cagtgaaaat gggcgacgat 1020
ttctatgccg ctacatatga tgagaaaaca ggtacaatta ctgctaaaac aaccacttat 1080
acagatggtg ctggcgttgc tcaaactgga gctgtgaaat ttggtggcgc aaatggtaaa 1140
tctgaagttg ttactgctac cgatggtaaa acttacttag caagcgacct tgacaaacat 1200
aacttcagaa caggcggtga gcttaaagag gttaatacag ataagactga aaacccactg 1260
cagaaaattg atgctgcctt ggcacaggtt gatacacttc gttctgacct gggtgcggta 1320
cagaaccgtt tcaactccgc tatcaccaac ctgggcaata ccgtaaataa cctgtcttct 1380
gcccgtagcc gtatcgaaga ttccgactac gcgaccgaag tctccaacat gtctcgcgcg 1440
cagattctgc agcaggccgg tacctccgtt ctggcgcagg ctaaccaggt tccgcaaaac 1500
gtcctctctt tactgcgtta a 1521

<210> 663
<211> 12
<212> PRT
<213> Artificial Sequence

<220>
<223> Linker

<400> 663
Glu Phe Ser Arg Tyr Pro Ala Gln Trp Arg Pro Leu
1 5 10

<210> 664
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Linker

<400> 664
gaattctcta gatatccagc acagtggcgg ccgctc 36

<210> 665
<211> 45
<212> PRT
<213> Artificial Sequence

<220>
<223> Linker

<400> 665
Lys Gly Asn Ser Lys Leu Glu Gly Gln Leu Glu Phe Pro Arg Thr Ser
1 5 10 15
Pro Val Trp Trp Asn Ser Ala Asp Ile Gln His Ser Gly Gly Arg Gln
20 25 30
Cys Asp Gly Tyr Leu Gln Asn Ser Pro Leu Arg Pro Leu
35 40 45

<210> 666
<211> 135
<212> DNA
<213> Artificial Sequence

<220>
<223> Linker

<400> 666
aagggcaatt cgaagcttga aggtcaattg gaattcccta ggactagtcc agtgtggtgg 60
aattctgcag atatccagca cagtggcggc cgccagtgtg atggatatct gcagaattcg 120
cccttgcggc cgctc 135

<210> 667
<211> 192
<212> PRT
<213> Artificial Sequence

<220>
<223> Hepatitis C E1

<400> 667
Tyr Gln Val Arg Asn Ser Thr Gly Leu Tyr His Val Thr Asn Asp Cys
1 5 10 15
Pro Asn Ser Ser Ile Val Tyr Glu Ala Ala Asp Ala Ile Leu His Thr
20 25 30
Pro Gly Cys Val Pro Cys Val Arg Glu Gly Asn Ala Ser Arg Cys Trp
35 40 45
Val Ala Met Thr Pro Thr Val Ala Thr Arg Asp Gly Lys Leu Pro Ala
50 55 60
Thr Gln Leu Arg Arg His Ile Asp Leu Leu Val Gly Ser Ala Thr Leu
65 70 75 80
Cys Ser Ala Leu Tyr Val Gly Asp Leu Cys Gly Ser Val Phe Leu Val
85 90 95
Gly Gln Leu Phe Thr Phe Ser Pro Arg Arg His Trp Thr Thr Gln Gly
100 105 110
Cys Asn Cys Ser Ile Tyr Pro Gly His Ile Thr Gly His Arg Met Ala
115 120 125
Trp Asp Met Met Met Asn Trp Ser Pro Thr Thr Ala Leu Val Met Ala
130 135 140
Gln Leu Leu Arg Ile Pro Gln Ala Ile Leu Asp Met Ile Ala Gly Ala
145 150 155 160
His Trp Gly Val Leu Ala Gly Ile Ala Tyr Phe Ser Met Val Gly Asn
165 170 175
Trp Ala Lys Val Leu Val Val Leu Leu Leu Phe Ala Gly Val Asp Ala
180 185 190

<210> 668
<211> 363
<212> PRT
<213> Artificial Sequence

<220>
<223> Hepatitis C E2

<400> 668
Glu Thr His Val Thr Gly Gly Ser Ala Gly His Thr Val Ser Gly Phe
1 5 10 15
Val Ser Leu Leu Ala Pro Gly Ala Lys Gln Asn Val Gln Leu Ile Asn
20 25 30
Thr Asn Gly Ser Trp His Leu Asn Ser Thr Ala Leu Asn Cys Asn Asp
35 40 45
Ser Leu Asn Thr Gly Trp Leu Ala Gly Leu Phe Tyr His His Lys Phe
50 55 60
Asn Ser Ser Gly Cys Pro Glu Arg Leu Ala Ser Cys Arg Pro Leu Thr
65 70 75 80
Asp Phe Asp Gln Gly Trp Gly Pro Ile Ser Tyr Ala Asn Gly Ser Gly
85 90 95
Pro Asp Gln Arg Pro Tyr Cys Trp His Tyr Pro Pro Lys Pro Cys Gly
100 105 110
Ile Val Pro Ala Lys Ser Val Cys Gly Pro Val Tyr Cys Phe Thr Pro
115 120 125
Ser Pro Val Val Val Gly Thr Thr Asp Arg Ser Gly Ala Pro Thr Tyr
130 135 140
Ser Trp Gly Glu Asn Asp Thr Asp Val Phe Val Leu Asn Asn Thr Arg
145 150 155 160
Pro Pro Leu Gly Asn Trp Phe Gly Cys Thr Trp Met Asn Ser Thr Gly
165 170 175
Phe Thr Lys Val Cys Gly Ala Pro Pro Cys Val Ile Gly Gly Ala Gly
180 185 190
Asn Asn Thr Leu His Cys Pro Thr Asp Cys Phe Arg Lys His Pro Asp
195 200 205
Ala Thr Tyr Ser Arg Cys Gly Ser Gly Pro Trp Ile Thr Pro Arg Cys
210 215 220
Leu Val Asp Tyr Pro Tyr Arg Leu Trp His Tyr Pro Cys Thr Ile Asn
225 230 235 240
Tyr Thr Ile Phe Lys Ile Arg Met Tyr Val Gly Gly Val Glu His Arg
245 250 255
Leu Glu Ala Ala Cys Asn Trp Thr Arg Gly Glu Arg Cys Asp Leu Glu
260 265 270
Asp Arg Asp Arg Ser Glu Leu Ser Pro Leu Leu Leu Thr Thr Thr Gln
275 280 285
Trp Gln Val Leu Pro Cys Ser Phe Thr Thr Leu Pro Ala Leu Ser Thr
290 295 300
Gly Leu Ile His Leu His Gln Asn Ile Val Asp Val Gln Tyr Leu Tyr
305 310 315 320
Gly Val Gly Ser Ser Ile Ala Ser Trp Ala Ile Lys Trp Glu Tyr Val
325 330 335
Val Leu Leu Phe Leu Leu Leu Ala Asp Ala Arg Val Cys Ser Cys Leu
340 345 350
Trp Met Met Leu Leu Ile Ser Gln Ala Glu Ala
355 360

<210> 669
<211> 576
<212> DNA
<213> Artificial Sequence

<220>
<223> Hepatitis C E1

<400> 669
taccaagtgc gcaactccac ggggctctac cacgtcacca atgattgccc taactcgagt 60
attgtgtacg aggcggccga tgccatcctg cacactccgg ggtgcgtccc ttgcgttcgc 120
gagggcaacg cctcgaggtg ttgggtggcg atgaccccta cggtggccac cagggatggc 180
aaactccccg cgacgcagct tcgacgtcac atcgatctgc ttgtcgggag cgccaccctc 240
tgttcggccc tctacgtggg ggacctgtgc gggtctgtct ttcttgtcgg ccaactgttt 300
accttctctc ccaggcgcca ctggacgacg caaggttgca attgctctat ctatcccggc 360
catataacgg gtcaccgcat ggcatgggat atgatgatga actggtcccc tacgacggcg 420
ttggtaatgg ctcagctgct ccggatccca caagccatct tggacatgat cgctggtgct 480
cactggggag tcctggcggg catagcgtat ttctccatgg tggggaactg ggcgaaggtc 540
ctggtagtgc tgctgctatt tgccggcgtc gacgcg 576

<210> 670
<211> 1089
<212> DNA
<213> Artificial Sequence

<220>
<223> Hepatitis C E2

<400> 670
gaaacccacg tcaccggggg aagtgccggc cacactgtgt ctggatttgt tagcctcctc 60
gcaccaggcg ccaagcagaa cgtccagctg atcaacacca acggcagttg gcacctcaat 120
agcacggccc tgaactgcaa tgatagcctc aacaccggct ggttggcagg gcttttctat 180
caccacaagt tcaactcttc aggctgtcct gagaggctag ccagctgccg acccccttacc 240
gattttgacc agggctgggg ccctatcagt tatgccaacg gaagcggccc cgaccagcgc 300
ccctactgct ggcactaccc cccaaaacct tgcggtattg tgcccgcgaa gagtgtgtgt 360
ggtccggtat attgcttcac tcccagcccc gtggtggtgg gaacgaccga caggtcgggc 420
gcgcccacct acagctgggg tgaaaatgat acggacgtct tcgtccttaa caataccagg 480
ccaccgctgg gcaattggtt cggttgtacc tggatgaact caactggatt caccaaagtg 540
tgcggagcgc ctccttgtgt catcggaggg gcgggcaaca acaccctgca ctgccccact 600
gattgcttcc gcaagcatcc ggacgccaca tactctcggt gcggctccgg tccctggatc 660
acacccaggt gcctggtcga ctacccgtat aggctttggc attatccttg taccatcaac 720
tacactatat ttaaaatcag gatgtacgtg ggaggggtcg agcacaggct ggaagctgcc 780
tgcaactgga cgcggggcga acgttgcgat ctggaagata gggacaggtc cgagctcagc 840
ccgttactgc tgaccactac acagtggcag gtcctcccgt gttccttcac aaccctgcca 900
gccttgtcca ccggcctcat ccacctccac cagaacattg tggacgtgca gtacttgtac 960
ggggtggggt caagcatcgc gtcctgggcc attaagtggg agtacgtcgt cctcctgttc 1020
cttctgcttg cagacgcgcg cgtctgctcc tgcttgtgga tgatgctact catatcccaa 1080
gcggaagcg 1089

<210> 671
<211> 595
<212> PRT
<213> E. coli

<400> 671
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Ile Thr Gln Asn
 1   5   10   15

```
Asn Ile Asn Lys Asn Gln Ser Ala Leu Ser Ser Ser Ile Glu Arg Leu
            20                  25                  30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
        35                  40                  45
Ala Ile Ala Asn Arg Phe Thr Ser Asn Ile Lys Gly Leu Thr Gln Ala
    50                  55                  60
Ala Arg Asn Ala Asn Asp Gly Ile Ser Val Ala Gln Thr Thr Glu Gly
65                  70                  75                  80
Ala Leu Ser Glu Ile Asn Asn Asn Leu Gln Arg Ile Arg Glu Leu Thr
                85                  90                  95
Val Gln Ala Ser Thr Gly Thr Asn Ser Asp Ser Asp Leu Asp Ser Ile
            100                 105                 110
Gln Asp Glu Ile Lys Ser Arg Leu Asp Glu Ile Asp Arg Val Ser Gly
        115                 120                 125
Gln Thr Gln Phe Asn Gly Val Asn Val Leu Ala Lys Asp Gly Ser Met
    130                 135                 140
Lys Ile Gln Val Gly Ala Asn Asp Gly Gln Thr Ile Thr Ile Asp Leu
145                 150                 155                 160
Lys Lys Ile Asp Ser Asp Thr Leu Gly Leu Asn Gly Phe Asn Val Asn
                165                 170                 175
Gly Ser Gly Thr Ile Ala Asn Lys Ala Ala Thr Ile Ser Asp Leu Thr
            180                 185                 190
Ala Ala Lys Met Asp Ala Ala Thr Asn Thr Ile Thr Thr Thr Asn Asn
        195                 200                 205
Ala Leu Thr Ala Ser Lys Ala Leu Asp Gln Leu Lys Asp Gly Asp Thr
    210                 215                 220
Val Thr Ile Lys Ala Asp Ala Ala Gln Thr Ala Thr Val Tyr Thr Tyr
225                 230                 235                 240
Asn Ala Ser Ala Gly Asn Phe Ser Phe Ser Asn Val Ser Asn Asn Thr
                245                 250                 255
Ser Ala Lys Ala Gly Asp Val Ala Ala Ser Leu Leu Pro Pro Ala Gly
            260                 265                 270
Gln Thr Ala Ser Gly Val Tyr Lys Ala Ala Ser Gly Glu Val Asn Phe
        275                 280                 285
Asp Val Asp Ala Asn Gly Lys Ile Thr Ile Gly Gly Gln Lys Ala Tyr
    290                 295                 300
Leu Thr Ser Asp Gly Asn Leu Thr Thr Asn Asp Ala Gly Gly Ala Thr
305                 310                 315                 320
Ala Ala Thr Leu Asp Gly Leu Phe Lys Lys Ala Gly Asp Gly Gln Ser
                325                 330                 335
Ile Gly Phe Lys Lys Thr Ala Ser Val Thr Met Gly Gly Thr Thr Tyr
            340                 345                 350
Asn Phe Lys Thr Gly Ala Asp Ala Asp Ala Ala Thr Ala Asn Ala Gly
        355                 360                 365
Val Ser Phe Thr Asp Thr Ala Ser Lys Glu Thr Val Leu Asn Lys Val
    370                 375                 380
Ala Thr Ala Lys Gln Gly Lys Ala Val Ala Ala Asp Gly Asp Thr Ser
385                 390                 395                 400
Ala Thr Ile Thr Tyr Lys Ser Gly Val Gln Thr Tyr Gln Ala Val Phe
                405                 410                 415
Ala Ala Gly Asp Gly Thr Ala Ser Ala Lys Tyr Ala Asp Lys Ala Asp
            420                 425                 430
Val Ser Asn Ala Thr Ala Thr Tyr Thr Asp Ala Asp Gly Glu Met Thr
        435                 440                 445
Thr Ile Gly Ser Tyr Thr Thr Lys Tyr Ser Ile Asp Ala Asn Asn Gly
    450                 455                 460
Lys Val Thr Val Asp Ser Gly Thr Gly Thr Gly Lys Tyr Ala Pro Lys
465                 470                 475                 480
Val Gly Ala Glu Val Tyr Val Ser Ala Asn Gly Thr Leu Thr Thr Asp
```

```
                485                 490                 495
Ala Thr Ser Glu Gly Thr Val Thr Lys Asp Pro Leu Lys Ala Leu Asp
            500                 505                 510
Glu Ala Ile Ser Ser Ile Asp Lys Phe Arg Ser Ser Leu Gly Ala Ile
        515                 520                 525
Gln Asn Arg Leu Asp Ser Ala Val Thr Asn Leu Asn Asn Thr Thr Thr
    530                 535                 540
Asn Leu Ser Glu Ala Gln Ser Arg Ile Gln Asp Ala Asp Tyr Ala Thr
545                 550                 555                 560
Glu Val Ser Asn Met Ser Lys Ala Gln Ile Ile Gln Gln Ala Gly Asn
                565                 570                 575
Ser Val Leu Ala Lys Ala Asn Gln Val Pro Gln Gln Val Leu Ser Leu
            580                 585                 590
Leu Gln Gly
        595
```

<210> 672
<211> 1788
<212> DNA
<213> E. coli

<400> 672
atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgcagcggg tcaggcgatt gctaaccgtt tcacctctaa cattaaaggc 180
ctgactcagg cggcccgtaa cgccaacgac ggtatctccg ttgcgcagac caccgaaggc 240
gcgctgtccg aaatcaacaa caacttacag cgtatccgtg aactgacggt tcaggcttct 300
accgggacta actccgattc ggatctggac tccattcagg acgaaatcaa atcccgtctg 360
gacgaaattg accgcgtatc tggccagacc cagttcaacg gcgtgaacgt actggcgaaa 420
gacggttcaa tgaaaattca ggttggtgcg aatgacggcc agactatcac gattgatctg 480
aagaaaattg actcagatac gctggggctg aatggtttta acgtgaatgg ttccggtacg 540
atagccaata aagcggcgac cattagcgac ctgacagcag cgaaaatgga tgctgcaact 600
aatactataa ctacaacaaa taatgcgctg actgcatcaa aggcgcttga tcaactgaaa 660
gatggtgaca ctgttactat caaagcagat gctgctcaaa ctgccacggt ttatacatac 720
aatgcatcag ctggtaactt ctcattcagt aatgtatcga ataatacttc agcaaaagca 780
ggtgatgtag cagctagcct tctcccgccg gctgggcaaa ctgctagtgg tgtttataaa 840
gcagcaagcg gtgaagtgaa ctttgatgtt gatgcgaatg gtaaaatcac aatcggagga 900
cagaaagcat atttaactag tgatggtaac ttaactacaa acgatgctgg tggtgcgact 960
gcggctacgc ttgatggttt attcaagaaa gctggtgatg gtcaatcaat cgggtttaag 1020
aagactgcat cagtcacgat gggggggaaca acttataact ttaaaacggg tgctgatgct 1080
gatgctgcaa ctgctaacgc aggggtatcg ttcactgata cagctagcaa agaaaccgtt 1140
ttaaataaag tggctacagc taaacaaggc aaagcagttg cagctgacgg tgatacatcc 1200
gcaacaatta cctataaatc tggcgttcag acgtatcagg ctgtatttgc cgcaggtgac 1260
ggtactgcta gcgcaaaata tgccgataaa gctgacgttt ctaatgcaac agcaacatac 1320
actgatgctg atggtgaaat gactacaatt ggttcataca ccacgaagta ttcaatcgat 1380
gctaacaacg gcaaggtaac tgttgattct ggaactggta cgggtaaata tgcgccgaaa 1440
gtaggggctg aagtatatgt tagtgctaat ggtactttaa caacagatgc aactagcgaa 1500
ggcacagtaa caaaagatcc actgaaagct ctggatgaag ctatcagctc catcgacaaa 1560
ttccgttctt ccctgggtgc tatccagaac cgtctggatt ccgcagtcac caacctgaac 1620
aacaccacta ccaacctgtc cgaagcgcag tcccgtattc aggacgccga ctatgcgacc 1680
gaagtgtcca acatgtcgaa agcgcagatc attcagcagg ccggtaactc cgtgctggca 1740
aaagccaacc aggtaccgca gcaggttctg tctctgctgc agggttaa 1788

<210> 673
<211> 1224
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2delta.EIII+

<400> 673
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcatgg agcgccggtg 540
gatcctgcta gcccatggac cgaaaacccg ctgcagaaaa ttgatgccgc gctggcgcag 600
gtggatgcgc tgcgctctga tctgggtgcg gtacaaaacc gtttcaactc tgctatcacc 660
aacctgggca ataccgtaaa caatctgtct gaagcgcgta gccgtatcga agattccgac 720
tacgcgaccg aagtttccaa catgtctcgc gcgcagattt tgcagcaggc cggtacttcc 780
gttctggcgc aggctaacca ggtcccgcag aacgtgctgt ctctgttacg tgaattctgc 840
agatatccag cacagtggcg gccgctcatg gaaaaattgc agttgaaggg aacaacctat 900
ggcgtctgtt caaaggcttt caagtttctt gggactcccg cagacacagg tcacggcact 960
gtggtgttgg aattgcagta cactggcacg gatggacctt gcaaagttcc tatctcgtca 1020
gtggcttcat tgaacgacct aacgccagtg ggcagattgg tcactgtcaa ccctttgtt 1080
tcagtggcca cggccaacgc taaggtcctg attgaattgg aaccaccctt tggagactca 1140
tacatagtgg tgggcagagg agaacaacag atcaatcacc attggcacaa gtctggaagc 1200
agcattggca aacccttaat aagc 1224

<210> 674
<211> 408
<212> PRT
<213> Artificial Sequence

<220>
<223> STF2delta.EIII+

<400> 674
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
1 5 10 15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
20 25 30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
35 40 45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
50 55 60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65 70 75 80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
85 90 95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
100 105 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
115 120 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
130 135 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145 150 155 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val His
165 170 175

```
Gly Ala Pro Val Asp Pro Ala Ser Pro Trp Thr Glu Asn Pro Leu Gln
            180                 185                 190
Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala Leu Arg Ser Asp Leu
        195                 200                 205
Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile Thr Asn Leu Gly Asn
    210                 215                 220
Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg Ile Glu Asp Ser Asp
225                 230                 235                 240
Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala Gln Ile Leu Gln Gln
                245                 250                 255
Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln Val Pro Gln Asn Val
            260                 265                 270
Leu Ser Leu Leu Arg Glu Phe Cys Arg Tyr Pro Ala Gln Trp Arg Pro
        275                 280                 285
Leu Met Glu Lys Leu Gln Leu Lys Gly Thr Thr Tyr Gly Val Cys Ser
    290                 295                 300
Lys Ala Phe Lys Phe Leu Gly Thr Pro Ala Asp Thr Gly His Gly Thr
305                 310                 315                 320
Val Val Leu Glu Leu Gln Tyr Thr Gly Thr Asp Gly Pro Cys Lys Val
                325                 330                 335
Pro Ile Ser Ser Val Ala Ser Leu Asn Asp Leu Thr Pro Val Gly Arg
            340                 345                 350
Leu Val Thr Val Asn Pro Phe Val Ser Val Ala Thr Ala Asn Ala Lys
        355                 360                 365
Val Leu Ile Glu Leu Glu Pro Pro Phe Gly Asp Ser Tyr Ile Val Val
    370                 375                 380
Gly Arg Gly Glu Gln Gln Ile Asn His His Trp His Lys Ser Gly Ser
385                 390                 395                 400
Ser Ile Gly Lys Pro Leu Ile Ser
                405
```

<210> 675
<211> 408
<212> PRT
<213> Artificial Sequence

<220>
<223> STF2delta.EIII+

<400> 675

```
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
 1               5                  10                  15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
            20                  25                  30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
        35                  40                  45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
    50                  55                  60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65                  70                  75                  80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
                85                  90                  95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
            100                 105                 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
        115                 120                 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
```

```
    130                 135                 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145                 150                 155                 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val His
                165                 170                 175
Gly Ala Pro Val Asp Pro Ala Ser Pro Trp Thr Glu Asn Pro Leu Gln
            180                 185                 190
Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala Leu Arg Ser Asp Leu
        195                 200                 205
Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile Thr Asn Leu Gly Asn
    210                 215                 220
Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg Ile Glu Asp Ser Asp
225                 230                 235                 240
Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala Gln Ile Leu Gln Gln
                245                 250                 255
Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln Val Pro Gln Asn Val
            260                 265                 270
Leu Ser Leu Leu Arg Glu Phe Ser Arg Tyr Pro Ala Gln Trp Arg Pro
        275                 280                 285
Leu Met Glu Lys Leu Gln Leu Lys Gly Thr Thr Tyr Gly Val Cys Ser
    290                 295                 300
Lys Ala Phe Lys Phe Leu Gly Thr Pro Ala Asp Thr Gly His Gly Thr
305                 310                 315                 320
Val Val Leu Glu Leu Gln Tyr Thr Gly Thr Asp Gly Pro Cys Lys Val
                325                 330                 335
Pro Ile Ser Ser Val Ala Ser Leu Asn Asp Leu Thr Pro Val Gly Arg
            340                 345                 350
Leu Val Thr Val Asn Pro Phe Val Ser Val Ala Thr Ala Asn Ala Lys
        355                 360                 365
Val Leu Ile Glu Leu Glu Pro Pro Phe Gly Asp Ser Tyr Ile Val Val
    370                 375                 380
Gly Arg Gly Glu Gln Gln Ile Asn His His Trp His Lys Ser Gly Ser
385                 390                 395                 400
Ser Ile Gly Lys Pro Leu Ile Ser
                405
```

<210> 676
<211> 1227
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2delta.EIII+

<400> 676
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcatgg agcgccggtg 540
gatcctgcta gcccatggac cgaaaacccg ctgcagaaaa ttgatgccgc gctggcgcag 600
gtggatgcgc tgcgctctga tctgggtgcg gtacaaaacc gtttcaactc tgctatcacc 660
aacctgggca ataccgtaaa caatctgtct gaagcgcgta gccgtatcga agattccgac 720

tacgcgaccg aagtttccaa catgtctcgc gcgcagattt tgcagcaggc cggtacttcc 780
gttctggcgc aggctaacca ggtcccgcag aacgtgctgt ctctgttacg tgaattctct 840
agatatccag cacagtggcg gccgctcatg gaaaaattgc agttgaaggg aacaacctat 900
ggcgtctgtt caaaggcttt caagtttctt gggactcccg cagacacagg tcacggcact 960
gtggtgttgg aattgcagta cactggcacg gatggacctt gcaaagttcc tatctcgtca 1020
gtggcttcat tgaacgacct aacgccagtg ggcagattgg tcactgtcaa ccctttgtt 1080
tcagtggcca cggccaacgc taaggtcctg attgaattgg aaccaccctt tggagactca 1140
tacatagtgg tgggcagagg agaacaacag atcaatcacc attggcacaa gtctggaagc 1200
agcattggca aacccttaat aagctga 1227

<210> 677
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<223> Linker

<400> 677
gaattctcta gatatccagc acagtggcgg ccgct 35

<210> 678
<211> 12
<212> PRT
<213> Artificial Sequence

<220>
<223> Linker

<400> 678
Glu Phe Ser Arg Tyr Pro Ala Gln Trp Arg Pro Leu
1 5 10

<210> 679
<211> 405
<212> PRT
<213> Artificial Sequence

<220>
<223> pET/STF2delta.JEIII+

<400> 679
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
1 5 10 15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
20 25 30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
35 40 45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
50 55 60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65 70 75 80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
85 90 95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
100 105 110

```
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
        115                 120                 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
    130                 135                 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145                 150                 155                 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val His
                165                 170                 175
Gly Ala Pro Val Asp Pro Ala Ser Pro Trp Thr Glu Asn Pro Leu Gln
            180                 185                 190
Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala Leu Arg Ser Asp Leu
        195                 200                 205
Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile Thr Asn Leu Gly Asn
    210                 215                 220
Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg Ile Glu Asp Ser Asp
225                 230                 235                 240
Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala Gln Ile Leu Gln Gln
                245                 250                 255
Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln Val Pro Gln Asn Val
            260                 265                 270
Leu Ser Leu Leu Arg Glu Phe Ser Arg Tyr Pro Ala Gln Trp Arg Pro
        275                 280                 285
Leu Met Asp Lys Leu Ala Leu Lys Gly Thr Thr Tyr Gly Met Cys Thr
    290                 295                 300
Glu Lys Phe Ser Phe Ala Lys Asn Pro Val Asp Thr Gly His Gly Thr
305                 310                 315                 320
Val Val Ile Glu Leu Ser Tyr Ser Gly Ser Asp Gly Pro Cys Lys Ile
                325                 330                 335
Pro Ile Val Ser Val Ala Ser Leu Asn Asp Met Thr Pro Val Gly Arg
            340                 345                 350
Leu Val Thr Val Asn Pro Phe Val Ala Thr Ser Ser Ala Asn Ser Lys
        355                 360                 365
Val Leu Val Glu Met Glu Pro Pro Phe Gly Asp Ser Tyr Ile Val Val
    370                 375                 380
Gly Arg Gly Asp Lys Gln Ile Asn His His Trp His Lys Ala Gly Ser
385                 390                 395                 400
Thr Leu Gly Lys Ala
                405
```

<210> 680
<211> 1215
<212> DNA
<213> Artificial Sequence

<220>
<223> pET/STF2delta.JEIII+

<400> 680
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcatgg agcgccggtg 540

gatcctgcta gcccatggac cgaaaacccg ctgcagaaaa ttgatgccgc gctggcgcag 600
gtggatgcgc tgcgctctga tctgggtgcg gtacaaaacc gtttcaactc tgctatcacc 660
aacctgggca ataccgtaaa caatctgtct gaagcgcgta gccgtatcga agattccgac 720
tacgcgaccg aagtttccaa catgtctcgc gcgcagattt tgcagcaggc cggtacttcc 780
gttctggcgc aggctaacca ggtcccgcag aacgtgctgt ctctgttacg tgaattcagc 840
agatatccag cacagtggcg gccgctcatg gacaaactgg ctctgaaagg cacaacctat 900
ggcatgtgta cagaaaaatt ctcgttcgcg aaaaatccgg tggacactgg tcacggaaca 960
gttgtcattg aactctccta ctctgggagt gatggcccct gcaaaattcc gattgtttcc 1020
gttgcgagcc tcaatgacat gacccccgtt gggcggctgg tgacagtgaa ccccttcgtc 1080
gcgacttcca gtgccaactc aaaggtgctg gtcgagatgg aacccccctt cggagactcc 1140
tacatcgtag ttggaagggg agacaagcag atcaaccacc attggcacaa agctggaagc 1200
acgctgggca aggcc 1215

<210> 681
<211> 348
<212> DNA
<213> Artificial Sequence

<220>
<223> JEIII+

<400> 681
atggacaaac tggctctgaa aggcacaacc tatggcatgt gtacagaaaa attctcgttc 60
gcgaaaaatc cggtggacac tggtcacgga acagttgtca ttgaactctc ctactctggg 120
agtgatggcc cctgcaaaat tccgattgtt tccgttgcga gcctcaatga catgacccca 180
gttgggcggc tggtgacagt gaacccttc gtcgcgactt ccagtgccaa ctcaaaggtg 240
ctggtcgaga tggaaccccc cttcggagac tcctacatcg tagttggaag gggagacaag 300
cagatcaacc accattggca caaagctgga agcacgctgg gcaaggcc 348

<210> 682
<211> 116
<212> PRT
<213> Artificial Sequence

<220>
<223> JEIII+

<400> 682

Met	Asp	Lys	Leu	Ala	Leu	Lys	Gly	Thr	Thr	Tyr	Gly	Met	Cys	Thr	Glu
1				5				10					15		
Lys	Phe	Ser	Phe	Ala	Lys	Asn	Pro	Val	Asp	Thr	Gly	His	Gly	Thr	Val
			20					25					30		
Val	Ile	Glu	Leu	Ser	Tyr	Ser	Gly	Ser	Asp	Gly	Pro	Cys	Lys	Ile	Pro
		35					40					45			
Ile	Val	Ser	Val	Ala	Ser	Leu	Asn	Asp	Met	Thr	Pro	Val	Gly	Arg	Leu
	50					55					60				
Val	Thr	Val	Asn	Pro	Phe	Val	Ala	Thr	Ser	Ser	Ala	Asn	Ser	Lys	Val
65					70					75					80
Leu	Val	Glu	Met	Glu	Pro	Pro	Phe	Gly	Asp	Ser	Tyr	Ile	Val	Val	Gly
				85					90					95	
Arg	Gly	Asp	Lys	Gln	Ile	Asn	His	His	Trp	His	Lys	Ala	Gly	Ser	Thr
			100					105					110		
Leu	Gly	Lys	Ala												
		115													

<210> 683

<211> 389
<212> PRT
<213> Artificial Sequence

<220>
<223> pET/STF2delta.Den1 EIII

<400> 683

```
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
 1               5                  10                  15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
            20                  25                  30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
        35                  40                  45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
    50                  55                  60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65                  70                  75                  80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
                85                  90                  95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
            100                 105                 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
        115                 120                 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
    130                 135                 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145                 150                 155                 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val His
                165                 170                 175
Gly Ala Pro Val Asp Pro Ala Ser Pro Trp Thr Glu Asn Pro Leu Gln
            180                 185                 190
Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala Leu Arg Ser Asp Leu
        195                 200                 205
Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile Thr Asn Leu Gly Asn
    210                 215                 220
Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg Ile Glu Asp Ser Asp
225                 230                 235                 240
Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala Gln Ile Leu Gln Gln
                245                 250                 255
Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln Val Pro Gln Asn Val
            260                 265                 270
Leu Ser Leu Leu Arg Glu Phe Ser Arg Tyr Pro Ala Gln Trp Arg Pro
        275                 280                 285
Leu Lys Gly Met Ser Tyr Val Met Cys Thr Gly Ser Phe Lys Leu Glu
    290                 295                 300
Lys Glu Val Ala Glu Thr Gln His Gly Thr Val Leu Val Gln Val Lys
305                 310                 315                 320
Tyr Glu Gly Thr Asp Ala Pro Cys Lys Ile Pro Phe Ser Thr Gln Asp
                325                 330                 335
Glu Lys Gly Val Thr Gln Asn Gly Arg Leu Ile Thr Ala Asn Pro Ile
            340                 345                 350
Val Thr Asp Lys Glu Lys Pro Val Asn Ile Glu Ala Glu Pro Pro Phe
        355                 360                 365
Gly Glu Asn Tyr Ile Val Val Gly Ala Gly Glu Lys Ala Leu Lys Leu
    370                 375                 380
Ser Trp Phe Lys Lys
385
```

<210> 684
<211> 1185
<212> DNA
<213> Artificial Sequence

<220>
<223> pET/STF2delta.Den1 EIII

<400> 684
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcatgg agcgccggtg 540
gatcctgcta gcccatggac cgaaaacccg ctgcagaaaa ttgatgccgc gctggcgcag 600
gtggatgcgc tgcgctctga tctgggtgcg gtacaaaacc gtttcaactc tgctatcacc 660
aacctgggca ataccgtaaa caatctgtct gaagcgcgta gccgtatcga agattccgac 720
tacgcgaccg aagtttccaa catgtctcgc gcgcagattt tgcagcaggc cggtacttcc 780
gttctggcgc aggctaacca ggtcccgcag aacgtgctgt ctctgttacg tgaattcagc 840
agatatccag cacagtggcg gccgctcaaa ggaatgtctt acgtaatgtg cacaggcagt 900
ttcaagctgg aaaaagaagt tgccgaaaca cagcatggca cggtactagt ccaagtgaaa 960
tatgagggaa cagacgcgcc atgtaagata ccattttcca ctcaagatga gaaaggggcg 1020
actcagaacg gaagattgat aaccgcaaat cctatcgtaa ccgacaagga aaagcccgtg 1080
aatattgagg cagagcctcc gtttggggag tcgtatatcg tcgttggtgc tggtgaaaag 1140
gctttaaagc tcagttggtt caaaaagggg tcaagcattg gtaaa 1185

<210> 685
<211> 389
<212> PRT
<213> Artificial Sequence

<220>
<223> pET/STF2delta.Den2 EIII

<400> 685
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
1 5 10 15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
20 25 30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
35 40 45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
50 55 60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65 70 75 80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
85 90 95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
100 105 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
115 120 125

```
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
    130                 135                 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145                 150                 155                 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val His
                165                 170                 175
Gly Ala Pro Val Asp Pro Ala Ser Pro Trp Thr Glu Asn Pro Leu Gln
            180                 185                 190
Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala Leu Arg Ser Asp Leu
        195                 200                 205
Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile Thr Asn Leu Gly Asn
    210                 215                 220
Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg Ile Glu Asp Ser Asp
225                 230                 235                 240
Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala Gln Ile Leu Gln Gln
                245                 250                 255
Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln Val Pro Gln Asn Val
            260                 265                 270
Leu Ser Leu Leu Arg Glu Phe Ser Arg Tyr Pro Ala Gln Trp Arg Pro
        275                 280                 285
Leu Lys Gly Met Ser Tyr Ser Met Cys Thr Gly Lys Phe Lys Val Val
    290                 295                 300
Lys Glu Ile Ala Glu Thr Gln His Gly Thr Ile Val Ile Arg Val Gln
305                 310                 315                 320
Tyr Glu Gly Asp Gly Ser Pro Cys Lys Thr Pro Phe Glu Ile Met Asp
                325                 330                 335
Leu Glu Lys Arg His Val Leu Gly Arg Leu Thr Thr Val Asn Pro Ile
            340                 345                 350
Val Thr Glu Lys Asp Ser Pro Val Asn Ile Glu Ala Glu Pro Pro Phe
        355                 360                 365
Gly Asp Ser Tyr Ile Ile Ile Gly Val Glu Pro Gly Gln Leu Lys Leu
    370                 375                 380
Asp Trp Phe Lys Lys
385
```

<210> 686
<211> 1185
<212> DNA
<213> Artificial Sequence

<220>
<223> pET/STF2delta.Den2 EIII

<400> 686
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcatgg agcgccggtg 540
gatcctgcta gcccatggac cgaaaacccg ctgcagaaaa ttgatgccgc gctggcgcag 600
gtggatgcgc tgcgctctga tctgggtgcg gtacaaaacc gtttcaactc tgctatcacc 660
aacctgggca ataccgtaaa caatctgtct gaagcgcgta gccgtatcga agattccgac 720
tacgcgaccg aagtttccaa catgtctcgc gcgcagattt tgcagcaggc cggtacttcc 780

gttctggcgc aggctaacca ggtcccgcag aacgtgctgt ctctgttacg tgaattcagc 840
agatatccag cacagtggcg gccgctcaaa ggtatgagct atagcatgtg taccggtaaa 900
tttaaagttg ttaaagaaat tgcggaaacc cagcatggta ccattgttat tcgtgttcag 960
tatgaaggtg atggtagccc gtgtaaaatt ccgtttgaaa ttatggatct ggaaaaacgt 1020
catgttctgg gtcgtctgat taccgttaat ccgattgtta ccgaaaaaga tagcccggtt 1080
aatattgaag cggaaccgcc gtttggtgat agctatatta ttattggtgt tgaaccgggt 1140
cagctgaaac tgaattggtt taaaaaaggt agcagcattg gtcag 1185

<210> 687
<211> 389
<212> PRT
<213> Artificial Sequence

<220>
<223> pET/STF2delta.Den3 EIII

<400> 687

```
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
 1               5                  10                  15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
            20                  25                  30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
        35                  40                  45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
    50                  55                  60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65                  70                  75                  80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
                85                  90                  95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
            100                 105                 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
        115                 120                 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
    130                 135                 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145                 150                 155                 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val His
                165                 170                 175
Gly Ala Pro Val Asp Pro Ala Ser Pro Trp Thr Glu Asn Pro Leu Gln
            180                 185                 190
Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala Leu Arg Ser Asp Leu
        195                 200                 205
Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile Thr Asn Leu Gly Asn
    210                 215                 220
Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg Ile Glu Asp Ser Asp
225                 230                 235                 240
Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala Gln Ile Leu Gln Gln
                245                 250                 255
Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln Val Pro Gln Asn Val
            260                 265                 270
Leu Ser Leu Leu Arg Glu Phe Ser Arg Tyr Pro Ala Gln Trp Arg Pro
        275                 280                 285
Leu Lys Gly Met Ser Tyr Ala Met Cys Leu Asn Thr Phe Val Leu Lys
    290                 295                 300
Lys Glu Val Ser Glu Thr Gln His Gly Thr Ile Leu Ile Lys Val Glu
305                 310                 315                 320
Tyr Lys Gly Glu Asp Ala Pro Cys Lys Ile Pro Phe Ser Thr Glu Asp
```

```
                325                 330                 335
Gly Gln Gly Lys Ala His Asn Gly Arg Leu Ile Thr Ala Asn Pro Val
            340                 345                 350
Val Thr Lys Lys Glu Glu Pro Val Asn Ile Glu Ala Glu Pro Pro Phe
        355                 360                 365
Gly Glu Ser Asn Ile Val Ile Gly Ile Gly Asp Lys Ala Leu Lys Ile
    370                 375                 380
Asn Trp Tyr Arg Lys
385
```

<210> 688
<211> 1185
<212> DNA
<213> Artificial Sequence

<220>
<223> pET/STF2delta.Den3 EIII

<400> 688
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcatgg agcgccggtg 540
gatcctgcta gcccatggac cgaaaacccg ctgcagaaaa ttgatgccgc gctggcgcag 600
gtggatgcgc tgcgctctga tctgggtgcg gtacaaaacc gtttcaactc tgctatcacc 660
aacctgggca ataccgtaaa caatctgtct gaagcgcgta gccgtatcga agattccgac 720
tacgcgaccg aagtttccaa catgtctcgc gcgcagattt tgcagcaggc cggtacttcc 780
gttctggcgc aggctaacca ggtcccgcag aacgtgctgt ctctgttacg tgaattcagc 840
agatatccag cacagtggcg gccgctcaaa ggaatgagtt atgcaatgtg tttaaataca 900
tttgtattaa aaaaagaagt aagtgaaaca caacatggaa caatattaat aaaagtagaa 960
tataaaggag aagatgcacc atgtaaaata ccatttagta cagaagatgg acaaggaaaa 1020
gcacataatg gaagattaat aacagcaaat ccagtagtaa caaaaaaaga agaaccagta 1080
aatatagaag cagaaccacc atttggagaa agtaatatag taataggaat aggagataaa 1140
gcattaaaaa taaattggta tagaaaagga agtagtatag gaaaa 1185

<210> 689
<211> 388
<212> PRT
<213> Artificial Sequence

<220>
<223> pET/STF2delta.Den4 EIII

<400> 689
```
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
 1               5                  10                  15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
            20                  25                  30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
        35                  40                  45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
```

```
    50                  55                  60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65                  70                  75                  80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
                85                  90                  95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
            100                 105                 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
        115                 120                 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
    130                 135                 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145                 150                 155                 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val His
                165                 170                 175
Gly Ala Pro Val Asp Pro Ala Ser Pro Trp Thr Glu Asn Pro Leu Gln
            180                 185                 190
Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala Leu Arg Ser Asp Leu
        195                 200                 205
Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile Thr Asn Leu Gly Asn
    210                 215                 220
Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg Ile Glu Asp Ser Asp
225                 230                 235                 240
Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala Gln Ile Leu Gln Gln
                245                 250                 255
Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln Val Pro Gln Asn Val
            260                 265                 270
Leu Ser Leu Leu Arg Glu Phe Ser Arg Tyr Pro Ala Gln Trp Arg Pro
        275                 280                 285
Leu Lys Gly Met Ser Tyr Thr Met Cys Ser Gly Lys Phe Ser Ile Asp
    290                 295                 300
Lys Glu Met Ala Glu Thr Gln His Gly Thr Thr Val Val Lys Val Lys
305                 310                 315                 320
Tyr Glu Gly Ala Gly Ala Pro Cys Lys Val Pro Ile Glu Ile Arg Asp
                325                 330                 335
Val Asn Lys Glu Lys Val Val Gly Arg Ile Ile Ser Pro Thr Pro Phe
            340                 345                 350
Ala Glu Asn Thr Asn Ser Val Thr Asn Ile Glu Leu Glu Arg Pro Leu
        355                 360                 365
Asp Ser Tyr Ile Val Ile Gly Val Gly Asp Ser Ala Leu Thr Leu His
    370                 375                 380
Trp Phe Arg Lys
385
```

<210> 690
<211> 1185
<212> DNA
<213> Artificial Sequence

<220>
<223> pET/STF2delta.Den4 EIII

<400> 690
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240

gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcatgg agcgccggtg 540
gatcctgcta gcccatggac cgaaaacccg ctgcagaaaa ttgatgccgc gctggcgcag 600
gtggatgcgc tgcgctctga tctgggtgcg gtacaaaacc gtttcaactc tgctatcacc 660
aacctgggca ataccgtaaa caatctgtct gaagcgcgta gccgtatcga agattccgac 720
tacgcgaccg aagtttccaa catgtctcgc gcgcagattt tgcagcaggc cggtacttcc 780
gttctggcgc aggctaacca ggtcccgcag aacgtgctgt ctctgttacg tgaattcagc 840
agatatccag cacagtggcg gccgctcaag ggaatgtcat acacgatgtg tagtggtaaa 900
ttctctatag acaaagagat ggcagagaca caacacggga caaccgtcgt gaaggttaag 960
tatgaaggag ctggcgcacc gtgcaaagta cccatcgaaa ttagggatgt aaacaaagag 1020
aaggtcgttg ggcgtatcat tagctcaacc ccacttgcgg aaaatactaa ttctgtaacg 1080
aacatagagt tggaaccacc ttttggtgat agctatatag ttattggtgt gggcaatagt 1140
gccttaactc tacattggtt tagaaaagga tcctcgatcg ggaaa 1185

<210> 691
<211> 27
<212> PRT
<213> Artificial Sequence

<220>
<223> West Nile virus

<400> 691
Leu Thr Ser Gly His Leu Lys Cys Arg Val Lys Met Glu Lys Leu Gln
1 5 10 15
Leu Lys Gly Thr Thr Tyr Gly Val Cys Ser Lys
20 25

<210> 692
<211> 27
<212> PRT
<213> Artificial Sequence

<220>
<223> Japanese encephalitis virus

<400> 692
Leu Thr Ser Gly His Leu Lys Cys Arg Leu Lys Met Asp Lys Leu Ala
1 5 10 15
Leu Lys Gly Thr Thr Tyr Gly Met Cys Thr Glu
20 25

<210> 693
<211> 27
<212> PRT
<213> Artificial Sequence

<220>
<223> Dengue virus

<400> 693

Ile Phe Ala Gly His Leu Lys Cys Arg Leu Lys Met Asp Lys Leu Thr
1 5 10 15
Leu Lys Gly Met Ser Tyr Val Met Cys Thr Gly
20 25

<210> 694
<211> 27
<212> PRT
<213> Artificial Sequence

<220>
<223> Dengue virus

<400> 694
Leu Phe Thr Gly His Leu Lys Cys Arg Leu Arg Met Asp Lys Leu Gln
1 5 10 15
Leu Lys Gly Met Ser Tyr Ser Met Cys Thr Gly
20 25

<210> 695
<211> 27
<212> PRT
<213> Artificial Sequence

<220>
<223> Dengue virus

<400> 695
Ile Phe Ala Gly His Leu Lys Cys Arg Leu Lys Met Asp Lys Leu Lys
1 5 10 15
Leu Lys Gly Met Ser Tyr Ala Met Cys Leu Asn
20 25

<210> 696
<211> 27
<212> PRT
<213> Artificial Sequence

<220>
<223> Dengue virus

<400> 696
Met Phe Ala Gly His Leu Lys Cys Lys Val Arg Met Glu Lys Leu Arg
1 5 10 15
Ile Lys Gly Met Ser Tyr Thr Met Cys Ser Gly
20 25

<210> 697
<211> 27
<212> PRT
<213> Artificial Sequence

<220>

<223> Flavivirus envelope protein

<220>
<221> VARIANT
<222> 1, 2, 3, 10, 11, 13, 16, 20, 21, 23, 24, 26, 27
<223> Xaa = Any Amino Acid

<400> 697
Xaa Xaa Xaa Gly His Leu Lys Cys Arg Xaa Xaa Met Xaa Lys Leu Xaa
1 5 10 15
Leu Lys Gly Xaa Xaa Tyr Xaa Xaa Cys Xaa Xaa
20 25

<210> 698
<211> 14
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 698
Gly His Leu Lys Cys Arg Met Lys Leu Leu Lys Gly Tyr Cys
1 5 10

<210> 699
<211> 100
<212> PRT
<213> Artificial Sequence

<220>
<223> Dengue virus

<400> 699
Lys Gly Met Ser Tyr Val Met Cys Thr Gly Ser Phe Lys Leu Glu Lys
1 5 10 15
Glu Val Ala Glu Thr Gln His Gly Thr Val Leu Val Gln Val Lys Tyr
20 25 30
Glu Gly Thr Asp Ala Pro Cys Lys Ile Pro Phe Ser Thr Gln Asp Glu
35 40 45
Lys Gly Val Thr Gln Asn Gly Arg Leu Ile Thr Ala Asn Pro Ile Val
50 55 60
Thr Asp Lys Glu Lys Pro Val Asn Ile Glu Ala Glu Pro Pro Phe Gly
65 70 75 80
Glu Asn Tyr Ile Val Val Gly Ala Gly Glu Lys Ala Leu Lys Leu Ser
85 90 95
Trp Phe Lys Lys
100

<210> 700
<211> 100
<212> PRT
<213> Artificial Sequence

<220>
<223> Dengue virus

<400> 700
Lys Gly Met Ser Tyr Ser Met Cys Thr Gly Lys Phe Lys Val Val Lys
1 5 10 15
Glu Ile Ala Glu Thr Gln His Gly Thr Ile Val Ile Arg Val Gln Tyr
20 25 30
Glu Gly Asp Gly Ser Pro Cys Lys Thr Pro Phe Glu Ile Met Asp Leu
35 40 45
Glu Lys Arg His Val Leu Gly Arg Leu Thr Thr Val Asn Pro Ile Val
50 55 60
Thr Glu Lys Asp Ser Pro Val Asn Ile Glu Ala Glu Pro Pro Phe Gly
65 70 75 80
Asp Ser Tyr Ile Ile Ile Gly Val Glu Pro Gly Gln Leu Lys Leu Asp
85 90 95
Trp Phe Lys Lys
100

<210> 701
<211> 100
<212> PRT
<213> Artificial Sequence

<220>
<223> Dengue virus

<400> 701
Lys Gly Met Ser Tyr Ala Met Cys Leu Asn Thr Phe Val Leu Lys Lys
1 5 10 15
Glu Val Ser Glu Thr Gln His Gly Thr Ile Leu Ile Lys Val Glu Tyr
20 25 30
Lys Gly Glu Asp Ala Pro Cys Lys Ile Pro Phe Ser Thr Glu Asp Gly
35 40 45
Gln Gly Lys Ala His Asn Gly Arg Leu Ile Thr Ala Asn Pro Val Val
50 55 60
Thr Lys Lys Glu Glu Pro Val Asn Ile Glu Ala Glu Pro Pro Phe Gly
65 70 75 80
Glu Ser Asn Ile Val Ile Gly Ile Gly Asp Lys Ala Leu Lys Ile Asn
85 90 95
Trp Tyr Arg Lys
100

<210> 702
<211> 99
<212> PRT
<213> Artificial Sequence

<220>
<223> Dengue virus

<400> 702
Lys Gly Met Ser Tyr Thr Met Cys Ser Gly Lys Phe Ser Ile Asp Lys

```
 1               5                  10                  15
Glu Met Ala Glu Thr Gln His Gly Thr Thr Val Val Lys Val Lys Tyr
            20                  25                  30
Glu Gly Ala Gly Ala Pro Cys Lys Val Pro Ile Glu Ile Arg Asp Val
        35                  40                  45
Asn Lys Glu Lys Val Val Gly Arg Ile Ile Ser Pro Thr Pro Phe Ala
    50                  55                  60
Glu Asn Thr Asn Ser Val Thr Asn Ile Glu Leu Glu Arg Pro Leu Asp
65                  70                  75                  80
Ser Tyr Ile Val Ile Gly Val Gly Asp Ser Ala Leu Thr Leu His Trp
                85                  90                  95
Phe Arg Lys
```

<210> 703
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 703
```
Cys Arg Val Lys Met Glu Lys Leu Gln Leu Lys Gly Thr Thr Tyr Gly
 1               5                  10                  15
Val Cys Ser Lys
            20
```

<210> 704
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 704
```
Ser Arg Val Lys Met Glu Lys Leu Gln Leu Lys Gly Thr Thr Tyr Gly
 1               5                  10                  15
Val Cys Ser Lys
            20
```

<210> 705
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 705
```
Cys Arg Val Lys Met Glu Lys Leu Gln Leu Lys Gly Thr Thr Tyr Gly
 1               5                  10                  15
```

Val Ser Ser Lys
20

<210> 706
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 706
Ser Arg Val Lys Met Glu Lys Leu Gln Leu Lys Gly Thr Thr Tyr Gly
1 5 10 15
Val Ser Ser Lys
20

<210> 707
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 707
Cys Arg Val Lys Met Glu Lys Leu Gln Leu Lys Gly Thr Thr Tyr Gly
1 5 10 15
Val Cys Ser Lys
20

<210> 708
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 708
Ala Arg Val Lys Met Glu Lys Leu Gln Leu Lys Gly Thr Thr Tyr Gly
1 5 10 15
Val Cys Ser Lys
20

<210> 709
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 709
Cys Ala Val Lys Met Glu Lys Leu Gln Leu Lys Gly Thr Thr Tyr Gly
1 5 10 15
Val Cys Ser Lys
20

<210> 710
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 710
Cys Arg Ala Lys Met Glu Lys Leu Gln Leu Lys Gly Thr Thr Tyr Gly
1 5 10 15
Val Cys Ser Lys
20

<210> 711
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 711
Cys Arg Val Ala Met Glu Lys Leu Gln Leu Lys Gly Thr Thr Tyr Gly
1 5 10 15
Val Cys Ser Lys
20

<210> 712
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 712
Cys Arg Val Lys Ala Glu Lys Leu Gln Leu Lys Gly Thr Thr Tyr Gly
1 5 10 15
Val Cys Ser Lys
20

<210> 713
<211> 20

<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 713
Cys Arg Val Lys Met Ala Lys Leu Gln Leu Lys Gly Thr Thr Tyr Gly
1 5 10 15
Val Cys Ser Lys
20

<210> 714
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 714
Cys Arg Val Lys Met Glu Ala Leu Gln Leu Lys Gly Thr Thr Tyr Gly
1 5 10 15
Val Cys Ser Lys
20

<210> 715
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 715
Cys Arg Val Lys Met Glu Lys Ala Gln Leu Lys Gly Thr Thr Tyr Gly
1 5 10 15
Val Cys Ser Lys
20

<210> 716
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 716
Cys Arg Val Lys Met Glu Lys Leu Ala Leu Lys Gly Thr Thr Tyr Gly
1 5 10 15
Val Cys Ser Lys

20

<210> 717
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 717
Cys Arg Val Lys Met Glu Lys Leu Gln Ala Lys Gly Thr Thr Tyr Gly
1 5 10 15
Val Cys Ser Lys
20

<210> 718
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 718
Cys Arg Val Lys Met Glu Lys Leu Gln Leu Ala Gly Thr Thr Tyr Gly
1 5 10 15
Val Cys Ser Lys
20

<210> 719
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 719
Cys Arg Val Lys Met Glu Lys Leu Gln Leu Lys Ala Thr Thr Tyr Gly
1 5 10 15
Val Cys Ser Lys
20

<210> 720
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 720
Cys Arg Val Lys Met Glu Lys Leu Gln Leu Lys Gly Ala Thr Tyr Gly
1 5 10 15
Val Cys Ser Lys
20

<210> 721
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 721
Cys Arg Val Lys Met Glu Lys Leu Gln Leu Lys Gly Thr Ala Tyr Gly
1 5 10 15
Val Cys Ser Lys
20

<210> 722
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 722
Cys Arg Val Lys Met Glu Lys Leu Gln Leu Lys Gly Thr Thr Ala Gly
1 5 10 15
Val Cys Ser Lys
20

<210> 723
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 723
Cys Arg Val Lys Met Glu Lys Leu Gln Leu Lys Gly Thr Thr Tyr Ala
1 5 10 15
Val Cys Ser Lys
20

<210> 724
<211> 20
<212> PRT

<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 724
Cys Arg Val Lys Met Glu Lys Leu Gln Leu Lys Gly Thr Thr Tyr Gly
1 5 10 15
Ala Cys Ser Lys
20

<210> 725
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 725
Cys Arg Val Lys Met Glu Lys Leu Gln Leu Lys Gly Thr Thr Tyr Gly
1 5 10 15
Val Ala Ser Lys
20

<210> 726
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 726
Cys Arg Val Lys Met Glu Lys Leu Gln Leu Lys Gly Thr Thr Tyr Gly
1 5 10 15
Val Cys Ala Lys
20

<210> 727
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 727
Cys Arg Val Lys Met Glu Lys Leu Gln Leu Lys Gly Thr Thr Tyr Gly
1 5 10 15
Val Cys Ser Ala
20

<210> 728
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 728
Cys Arg Val Lys Met Glu Lys Leu Gln Leu Lys Gly Thr Thr Tyr Gly
1 5 10 15
Val Cys Ser Lys
20

<210> 729
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 729
Ser Asn Thr Val Lys Leu Thr Ser Gly His Leu Lys Cys Arg Val Lys
1 5 10 15
Met Glu Lys Leu
20

<210> 730
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 730
Asn Thr Val Lys Leu Thr Ser Gly His Leu Lys Cys Arg Val Lys Met
1 5 10 15
Glu Lys Leu Gln
20

<210> 731
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 731
Thr Val Lys Leu Thr Ser Gly His Leu Lys Cys Arg Val Lys Met Glu
1 5 10 15
Lys Leu Gln Leu
20

<210> 732
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 732
Val Lys Leu Thr Ser Gly His Leu Lys Cys Arg Val Lys Met Glu Lys
1 5 10 15
Leu Gln Leu Lys
20

<210> 733
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 733
Lys Leu Thr Ser Gly His Leu Lys Cys Arg Val Lys Met Glu Lys Leu
1 5 10 15
Gln Leu Lys Gly
20

<210> 734
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 734
Leu Thr Ser Gly His Leu Lys Cys Arg Val Lys Met Glu Lys Leu Gln
1 5 10 15
Leu Lys Gly Thr
20

<210> 735
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 735
Thr Ser Gly His Leu Lys Cys Arg Val Lys Met Glu Lys Leu Gln Leu
1 5 10 15
Lys Gly Thr Thr
20

<210> 736
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 736
Ser Gly His Leu Lys Cys Arg Val Lys Met Glu Lys Leu Gln Leu Lys
1 5 10 15
Gly Thr Thr Tyr
20

<210> 737
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 737
Gly His Leu Lys Cys Arg Val Lys Met Glu Lys Leu Gln Leu Lys Gly
1 5 10 15
Thr Thr Tyr Gly
20

<210> 738
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 738
His Leu Lys Cys Arg Val Lys Met Glu Lys Leu Gln Leu Lys Gly Thr
1 5 10 15
Thr Tyr Gly Val
20

<210> 739
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 739
Leu Lys Cys Arg Val Lys Met Glu Lys Leu Gln Leu Lys Gly Thr Thr
1 5 10 15
Tyr Gly Val Cys
20

<210> 740
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 740
Lys Cys Arg Val Lys Met Glu Lys Leu Gln Leu Lys Gly Thr Thr Tyr
1 5 10 15
Gly Val Cys Ser
20

<210> 741
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 741
Cys Arg Val Lys Met Glu Lys Leu Gln Leu Lys Gly Thr Thr Tyr Gly
1 5 10 15
Val Cys Ser Lys
20

<210> 742
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 742
Arg Val Lys Met Glu Lys Leu Gln Leu Lys Gly Thr Thr Tyr Gly Val

1 5 10 15
Cys Ser Lys Ala
20

<210> 743
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 743
Val Lys Met Glu Lys Leu Gln Leu Lys Gly Thr Thr Tyr Gly Val Cys
1 5 10 15
Ser Lys Ala Phe
20

<210> 744
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 744
Lys Met Glu Lys Leu Gln Leu Lys Gly Thr Thr Tyr Gly Val Cys Ser
1 5 10 15
Lys Ala Phe Lys
20

<210> 745
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 745
Met Glu Lys Leu Gln Leu Lys Gly Thr Thr Tyr Gly Val Cys Ser Lys
1 5 10 15
Ala Phe Lys Phe
20

<210> 746
<211> 20
<212> PRT
<213> Artificial Sequence

<220>

<223> Flavivirus envelope protein

<400> 746
Glu Lys Leu Gln Leu Lys Gly Thr Thr Tyr Gly Val Cys Ser Lys Ala
1 5 10 15
Phe Lys Phe Leu
20

<210> 747
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 747
Lys Leu Gln Leu Lys Gly Thr Thr Tyr Gly Val Cys Ser Lys Ala Phe
1 5 10 15
Lys Phe Leu Gly
20

<210> 748
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 748
Leu Gln Leu Lys Gly Thr Thr Tyr Gly Val Cys Ser Lys Ala Phe Lys
1 5 10 15
Phe Leu Gly Thr
20

<210> 749
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 749
Gln Leu Lys Gly Thr Thr Tyr Gly Val Cys Ser Lys Ala Phe Lys Phe
1 5 10 15
Leu Gly Thr Pro
20

<210> 750

<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 750
Leu Lys Gly Thr Thr Tyr Gly Val Cys Ser Lys Ala Phe Lys Phe Leu
1 5 10 15
Gly Thr Pro Ala
20

<210> 751
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 751
Lys Gly Thr Thr Tyr Gly Val Cys Ser Lys Ala Phe Lys Phe Leu Gly
1 5 10 15
Thr Pro Ala Asp
20

<210> 752
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 752
Gly Thr Thr Tyr Gly Val Cys Ser Lys Ala Phe Lys Phe Leu Gly Thr
1 5 10 15
Pro Ala Asp Thr
20

<210> 753
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 753
Thr Thr Tyr Gly Val Cys Ser Lys Ala Phe Lys Phe Leu Gly Thr Pro
1 5 10 15

Ala Asp Thr Gly
20

<210> 754
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Flavivirus envelope protein

<400> 754
Thr Tyr Gly Val Cys Ser Lys Ala Phe Lys Phe Leu Gly Thr Pro Ala
1 5 10 15
Asp Thr Gly His
20

<210> 755
<211> 2760
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2.OVA

<400> 755
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcagaa agcgtatgat 540
gtgaaagata cagcagtaac aacgaaagct tatgccaata atggtactac actggatgta 600
tcgggtcttg atgatgcagc tattaaagcg gctacgggtg gtacgaatgg tacggcttct 660
gtaaccggtg gtgcggttaa atttgacgca gataataaca agtactttgt tactattggt 720
ggctttactg gtgctgatgc cgccaaaaat ggcgattatg aagttaacgt tgctactgac 780
ggtacagtaa cccttgcggc tggcgcaact aaaaccacaa tgcctgctgg tgcgacaact 840
aaaacagaag tacaggagtt aaaagataca ccggcagttg tttcagcaga tgctaaaaat 900
gccttaattg ctggcggcgt tgacgctacc gatgctaatg gcgctgagtt ggtcaaaatg 960
tcttataccg ataaaaatgg taagacaatt gaaggcggtt atgcgcttaa agctggcgat 1020
aagtattacg ccgcagatta cgatgaagcg acaggagcaa ttaaagctaa aaccacaagt 1080
tatactgctg ctgacggcac taccaaaaca gcggctaacc aactgggtgg cgtagacggt 1140
aaaaccgaag tcgttactat cgacggtaaa acctacaatg ccagcaaagc cgctggtcat 1200
gatttcaaag cacaaccaga gctggcggaa gcagccgcta aaaccaccga aaacccgctg 1260
cagaaaattg atgccgcgct ggcgcaggtg gatgcgctgc gctctgatct gggtgcggta 1320
caaaaccgtt tcaactctgc tatcaccaac ctgggcaata ccgtaaacaa tctgtctgaa 1380
gcgcgtagcc gtatcgaaga ttccgactac gcgaccgaag tttccaacat gtctcgcgcg 1440
cagattttgc agcaggccgg tacttccgtt ctggcgcagg ctaaccaggt cccgcagaac 1500
gtgctgtctc tgttacgtct cgagggctcc atcggcgcag caagcatgga attttgtttt 1560
gatgtattca aggagctcaa agtccaccat gccaatgaga acatcttcta ctgccccatt 1620
gccatcatgt cagctctagc catggtatac ctgggtgcaa aagacagcac caggacacaa 1680
ataaataagg ttgttcgctt tgataaactt ccaggattcg gagacagtat tgaagctcag 1740

```
tgtggcacat ctgtaaacgt tcactcttca cttagagaca tcctcaacca aatcaccaaa 1800
ccaaatgatg tttattcgtt cagccttgcc agtagacttt atgctgaaga gagataccca 1860
atcctgccag aatacttgca gtgtgtgaag gaactgtata gaggaggctt ggaacctatc 1920
aactttcaaa cagctgcaga tcaagccaga gagctcatca attcctgggt agaaagtcag 1980
acaaatggaa ttatcagaaa tgtccttcag ccaagctccg tggattctca aactgcaatg 2040
gttctggtta atgccattgt cttcaaagga ctgtgggaga aagcatttaa ggatgaagac 2100
acacaagcaa tgcctttcag agtgactgag caagaaagca aacctgtgca gatgatgtac 2160
cagattggtt tatttagagt ggcatcaatg gcttctgaga aaatgaagat cctggagctt 2220
ccatttgcca gtgggacaat gagcatgttg gtgctgttgc ctgatgaagt ctcaggcctt 2280
gagcagcttg agagtataat caactttgaa aaactgactg aatggaccag ttctaatgtt 2340
atggaagaga ggaagatcaa agtgtactta cctcgcatga agatggagga aaaatacaac 2400
ctcacatctg tcttaatggc tatgggcatt actgacgtgt ttagctcttc agccaatctg 2460
tctggcatct cctcagcaga gagcctgaag atatctcaag ctgtccatgc agcacatgca 2520
gaaatcaatg aagcaggcag agaggtggta gggtcagcag aggctggagt ggatgctgca 2580
agcgtctctg aagaatttag ggctgaccat ccattcctct tctgtatcaa gcacatcgca 2640
accaacgccg ttctcttctt tggcagatgt gtttcccctt cgaagcttga aggtaagcct 2700
atccctaacc ctctcctcgg tctcgattct acgcgtaccg gtcatcatca ccatcaccat 2760
```

<210> 756
<211> 920
<212> PRT
<213> Artificial Sequence

<220>
<223> STF2.OVA

<400> 756

```
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
 1               5                  10                  15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
            20                  25                  30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
        35                  40                  45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
    50                  55                  60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65                  70                  75                  80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
                85                  90                  95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
            100                 105                 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
        115                 120                 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
    130                 135                 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145                 150                 155                 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val Gln
                165                 170                 175
Lys Ala Tyr Asp Val Lys Asp Thr Ala Val Thr Thr Lys Ala Tyr Ala
            180                 185                 190
Asn Asn Gly Thr Thr Leu Asp Val Ser Gly Leu Asp Asp Ala Ala Ile
        195                 200                 205
Lys Ala Ala Thr Gly Gly Thr Asn Gly Thr Ala Ser Val Thr Gly Gly
    210                 215                 220
Ala Val Lys Phe Asp Ala Asp Asn Asn Lys Tyr Phe Val Thr Ile Gly
225                 230                 235                 240
```

```
Gly Phe Thr Gly Ala Asp Ala Ala Lys Asn Gly Asp Tyr Glu Val Asn
                245                 250                 255
Val Ala Thr Asp Gly Thr Val Thr Leu Ala Ala Gly Ala Thr Lys Thr
            260                 265                 270
Thr Met Pro Ala Gly Ala Thr Thr Lys Thr Glu Val Gln Glu Leu Lys
        275                 280                 285
Asp Thr Pro Ala Val Val Ser Ala Asp Ala Lys Asn Ala Leu Ile Ala
    290                 295                 300
Gly Gly Val Asp Ala Thr Asp Ala Asn Gly Ala Glu Leu Val Lys Met
305                 310                 315                 320
Ser Tyr Thr Asp Lys Asn Gly Lys Thr Ile Glu Gly Gly Tyr Ala Leu
                325                 330                 335
Lys Ala Gly Asp Lys Tyr Tyr Ala Ala Asp Tyr Asp Glu Ala Thr Gly
            340                 345                 350
Ala Ile Lys Ala Lys Thr Thr Ser Tyr Thr Ala Ala Asp Gly Thr Thr
        355                 360                 365
Lys Thr Ala Ala Asn Gln Leu Gly Gly Val Asp Gly Lys Thr Glu Val
    370                 375                 380
Val Thr Ile Asp Gly Lys Thr Tyr Asn Ala Ser Lys Ala Ala Gly His
385                 390                 395                 400
Asp Phe Lys Ala Gln Pro Glu Leu Ala Glu Ala Ala Ala Lys Thr Thr
                405                 410                 415
Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala
            420                 425                 430
Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile
        435                 440                 445
Thr Asn Leu Gly Asn Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg
    450                 455                 460
Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala
465                 470                 475                 480
Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln
                485                 490                 495
Val Pro Gln Asn Val Leu Ser Leu Leu Arg Leu Glu Gly Ser Ile Gly
            500                 505                 510
Ala Ala Ser Met Glu Phe Cys Phe Asp Val Phe Lys Glu Leu Lys Val
        515                 520                 525
His His Ala Asn Glu Asn Ile Phe Tyr Cys Pro Ile Ala Ile Met Ser
    530                 535                 540
Ala Leu Ala Met Val Tyr Leu Gly Ala Lys Asp Ser Thr Arg Thr Gln
545                 550                 555                 560
Ile Asn Lys Val Val Arg Phe Asp Lys Leu Pro Gly Phe Gly Asp Ser
                565                 570                 575
Ile Glu Ala Gln Cys Gly Thr Ser Val Asn Val His Ser Ser Leu Arg
            580                 585                 590
Asp Ile Leu Asn Gln Ile Thr Lys Pro Asn Asp Val Tyr Ser Phe Ser
        595                 600                 605
Leu Ala Ser Arg Leu Tyr Ala Glu Glu Arg Tyr Pro Ile Leu Pro Glu
    610                 615                 620
Tyr Leu Gln Cys Val Lys Glu Leu Tyr Arg Gly Gly Leu Glu Pro Ile
625                 630                 635                 640
Asn Phe Gln Thr Ala Ala Asp Gln Ala Arg Glu Leu Ile Asn Ser Trp
                645                 650                 655
Val Glu Ser Gln Thr Asn Gly Ile Ile Arg Asn Val Leu Gln Pro Ser
            660                 665                 670
Ser Val Asp Ser Gln Thr Ala Met Val Leu Val Asn Ala Ile Val Phe
        675                 680                 685
Lys Gly Leu Trp Glu Lys Ala Phe Lys Asp Glu Asp Thr Gln Ala Met
    690                 695                 700
Pro Phe Arg Val Thr Glu Gln Glu Ser Lys Pro Val Gln Met Met Tyr
```

705 710 715 720
Gln Ile Gly Leu Phe Arg Val Ala Ser Met Ala Ser Glu Lys Met Lys
725 730 735
Ile Leu Glu Leu Pro Phe Ala Ser Gly Thr Met Ser Met Leu Val Leu
740 745 750
Leu Pro Asp Glu Val Ser Gly Leu Glu Gln Leu Glu Ser Ile Ile Asn
755 760 765
Phe Glu Lys Leu Thr Glu Trp Thr Ser Ser Asn Val Met Glu Glu Arg
770 775 780
Lys Ile Lys Val Tyr Leu Pro Arg Met Lys Met Glu Glu Lys Tyr Asn
785 790 795 800
Leu Thr Ser Val Leu Met Ala Met Gly Ile Thr Asp Val Phe Ser Ser
805 810 815
Ser Ala Asn Leu Ser Gly Ile Ser Ser Ala Glu Ser Leu Lys Ile Ser
820 825 830
Gln Ala Val His Ala Ala His Ala Glu Ile Asn Glu Ala Gly Arg Glu
835 840 845
Val Val Gly Ser Ala Glu Ala Gly Val Asp Ala Ala Ser Val Ser Glu
850 855 860
Glu Phe Arg Ala Asp His Pro Phe Leu Phe Cys Ile Lys His Ile Ala
865 870 875 880
Thr Asn Ala Val Leu Phe Phe Gly Arg Cys Val Ser Pro Ser Lys Leu
885 890 895
Glu Gly Lys Pro Ile Pro Asn Pro Leu Leu Gly Leu Asp Ser Thr Arg
900 905 910
Thr Gly His His His His His His
915 920

<210> 757
<211> 385
<212> PRT
<213> Artificial Sequence

<220>
<223> OVA

<400> 757
Gly Ser Ile Gly Ala Ala Ser Met Glu Phe Cys Phe Asp Val Phe Lys
1 5 10 15
Glu Leu Lys Val His His Ala Asn Glu Asn Ile Phe Tyr Cys Pro Ile
20 25 30
Ala Ile Met Ser Ala Leu Ala Met Val Tyr Leu Gly Ala Lys Asp Ser
35 40 45
Thr Arg Thr Gln Ile Asn Lys Val Val Arg Phe Asp Lys Leu Pro Gly
50 55 60
Phe Gly Asp Ser Ile Glu Ala Gln Cys Gly Thr Ser Val Asn Val His
65 70 75 80
Ser Ser Leu Arg Asp Ile Leu Asn Gln Ile Thr Lys Pro Asn Asp Val
85 90 95
Tyr Ser Phe Ser Leu Ala Ser Arg Leu Tyr Ala Glu Glu Arg Tyr Pro
100 105 110
Ile Leu Pro Glu Tyr Leu Gln Cys Val Lys Glu Leu Tyr Arg Gly Gly
115 120 125
Leu Glu Pro Ile Asn Phe Gln Thr Ala Ala Asp Gln Ala Arg Glu Leu
130 135 140
Ile Asn Ser Trp Val Glu Ser Gln Thr Asn Gly Ile Ile Arg Asn Val
145 150 155 160

```
Leu Gln Pro Ser Ser Val Asp Ser Gln Thr Ala Met Val Leu Val Asn
                165                 170                 175
Ala Ile Val Phe Lys Gly Leu Trp Glu Lys Ala Phe Lys Asp Glu Asp
            180                 185                 190
Thr Gln Ala Met Pro Phe Arg Val Thr Glu Gln Glu Ser Lys Pro Val
        195                 200                 205
Gln Met Met Tyr Gln Ile Gly Leu Phe Arg Val Ala Ser Met Ala Ser
    210                 215                 220
Glu Lys Met Lys Ile Leu Glu Leu Pro Phe Ala Ser Gly Thr Met Ser
225                 230                 235                 240
Met Leu Val Leu Leu Pro Asp Glu Val Ser Gly Leu Glu Gln Leu Glu
                245                 250                 255
Ser Ile Ile Asn Phe Glu Lys Leu Thr Glu Trp Thr Ser Ser Asn Val
            260                 265                 270
Met Glu Glu Arg Lys Ile Lys Val Tyr Leu Pro Arg Met Lys Met Glu
        275                 280                 285
Glu Lys Tyr Asn Leu Thr Ser Val Leu Met Ala Met Gly Ile Thr Asp
    290                 295                 300
Val Phe Ser Ser Ser Ala Asn Leu Ser Gly Ile Ser Ser Ala Glu Ser
305                 310                 315                 320
Leu Lys Ile Ser Gln Ala Val His Ala Ala His Ala Glu Ile Asn Glu
                325                 330                 335
Ala Gly Arg Glu Val Val Gly Ser Ala Glu Ala Gly Val Asp Ala Ala
            340                 345                 350
Ser Val Ser Glu Glu Phe Arg Ala Asp His Pro Phe Leu Phe Cys Ile
        355                 360                 365
Lys His Ile Ala Thr Asn Ala Val Leu Phe Phe Gly Arg Cys Val Ser
    370                 375                 380
Pro
385
```

<210> 758
<211> 1156
<212> DNA
<213> Artificial Sequence

<220>
<223> OVA

<400> 758

```
gggctccatc ggcgcagcaa gcatggaatt ttgttttgat gtattcaagg agctcaaagt 60
ccaccatgcc aatgagaaca tcttctactg ccccattgcc atcatgtcag ctctagccat 120
ggtatacctg ggtgcaaaag acagcaccag gacacaaata ataaggttg ttcgctttga 180
taaacttcca ggattcggag acagtattga agctcagtgt ggcacatctg taaacgttca 240
ctcttcactt agagacatcc tcaaccaaat caccaaacca aatgatgttt attcgttcag 300
ccttgccagt agactttatg ctgaagagag atacccaatc ctgccagaat acttgcagtg 360
tgtgaaggaa ctgtatagag gaggcttgga acctatcaac tttcaaacag ctgcagatca 420
agccagagag ctcatcaatt cctgggtaga aagtcagaca aatggaatta tcagaaatgt 480
ccttcagcca agctccgtgg attctcaaac tgcaatggtt ctggttaatg ccattgtctt 540
caaaggactg tgggagaaag catttaagga tgaagacaca caagcaatgc ctttcagagt 600
gactgagcaa gaaagcaaac ctgtgcagat gatgtaccag attggtttat ttagagtggc 660
atcaatggct tctgagaaaa tgaagatcct ggagcttcca tttgccagtg ggacaatgag 720
catgttggtg ctgttgcctg atgaagtctc aggccttgag cagcttgaga gtataatcaa 780
ctttgaaaaa ctgactgaat ggaccagttc taatgttatg gaagagagga agatcaaagt 840
gtacttacct cgcatgaaga tggaggaaaa atacaacctc acatctgtct taatggctat 900
gggcattact gacgtgttta gctcttcagc caatctgtct ggcatctcct cagcagagag 960
cctgaagata tctcaagctg tccatgcagc acatgcagaa atcaatgaag caggcagaga 1020
```

ggtggtaggg tcagcagagg ctggagtgga tgctgcaagc gtctctgaag aatttagggc 1080
tgaccatcca ttcctcttct gtatcaagca catcgcaacc aacgccgttc tcttctttgg 1140
cagatgtgtt tcccct 1156

<210> 759
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 759
catctcagtg caactaaa 18

<210> 760
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 760
tagaaggcac agtcgagg 18

<210> 761
<211> 2736
<212> DNA
<213> Artificial Sequence

<220>
<223> STF2.E

<400> 761
atggcacaag taatcaacac taacagtctg tcgctgctga cccagaataa cctgaacaaa 60
tcccagtccg cactgggcac cgctatcgag cgtctgtctt ctggtctgcg tatcaacagc 120
gcgaaagacg atgcggcagg tcaggcgatt gctaaccgtt tcaccgcgaa catcaaaggt 180
ctgactcagg cttcccgtaa cgctaacgac ggtatctcca ttgcgcagac cactgaaggc 240
gcgctgaacg aaatcaacaa caacctgcag cgtgtgcgtg aactggcggt tcagtctgct 300
aacagcacca actcccagtc tgacctcgac tccatccagg ctgaaatcac ccagcgcctg 360
aacgaaatcg accgtgtatc cggccagact cagttcaacg gcgtgaaagt cctggcgcag 420
gacaacaccc tgaccatcca ggttggcgcc aacgacggtg aaactatcga tatcgatctg 480
aagcagatca actctcagac cctgggtctg gactcactga acgtgcagaa agcgtatgat 540
gtgaaagata cagcagtaac aacgaaagct tatgccaata atggtactac actggatgta 600
tcgggtcttg atgatgcagc tattaaagcg gctacgggtg gtacgaatgg tacggcttct 660
gtaaccggtg gtgcggttaa atttgacgca gataataaca agtactttgt tactattggt 720
ggctttactg gtgctgatgc cgccaaaaat ggcgattatg aagttaacgt tgctactgac 780
ggtacagtaa cccttgcggc tggcgcaact aaaaccacaa tgcctgctgg tgcgacaact 840
aaaacagaag tacaggagtt aaaagataca ccggcagttg tttcagcaga tgctaaaaat 900
gccttaattg ctggcggcgt tgacgctacc gatgctaatg gcgctgagtt ggtcaaaatg 960
tcttataccg ataaaaatgg taagacaatt gaaggcggtt atgcgcttaa agctggcgat 1020
aagtattacg ccgcagatta cgatgaagcg acaggagcaa ttaaagctaa aaccacaagt 1080
tatactgctg ctgacggcac taccaaaaca gcggctaacc aactgggtgg cgtagacggt 1140
aaaaccgaag tcgttactat cgacggtaaa acctacaatg ccagcaaagc cgctggtcat 1200
gatttcaaag cacaaccaga gctggcggaa gcagccgcta aaaccaccga aaacccgctg 1260

cagaaaattg atgccgcgct ggcgcaggtg gatgcgctgc gctctgatct gggtgcggta 1320
caaaaccgtt tcaactctgc tatcaccaac ctgggcaata ccgtaaacaa tctgtctgaa 1380
gcgcgtagcc gtatcgaaga ttccgactac gcgaccgaag tttccaacat gtctcgcgcg 1440
cagattttgc agcaggccgg tacttccgtt ctggcgcagg ctaaccaggt cccgcagaac 1500
gtgctgtctc tgttacgttt caactgcctt ggaatgagca acagagactt cttggaagga 1560
gtgtctggag caacatgggt ggatttggtt ctcgaaggcg acagctgcgt gactatcatg 1620
tctaaggaca agcctaccat cgatgtgaag atgatgaata tggaggcggc caacctggca 1680
gaggtccgca gttattgcta tttggctacc gtcagcgatc tctccaccaa agctgcgtgc 1740
ccgaccatgg gagaagctca caatgacaaa cgtgctgacc cagcttttgt gtgcagacaa 1800
ggagtggtgg acaggggctg gggcaacggc tgcggactat ttggcaaagg aagcattgac 1860
acatgcgcca aatttgcctg ctctaccaag gcaataggaa gaaccatctt gaaagagaat 1920
atcaagtacg aagtggccat ttttgtccat ggaccaacta ctgtggagtc gcacggaaac 1980
tactccacac aggttggagc cactcaggca gggagattca gcatcactcc tgcagcgcct 2040
tcatacacac taaagcttgg agaatatgga gaggtgacag tggactgtga accacggtca 2100
gggattgaca ccaatgcata ctacgtgatg actgttggaa caaagacgtt cttggtccat 2160
cgtgagtggt tcatggacct caacctccct tggagcagtg ctggaagtac tgtgtggagg 2220
aacagagaga cgttaatgga gtttgaggaa ccacacgcca cgaagcagtc tgtgatagca 2280
ttgggctcac aagagggagc tctgcatcaa gctttggctg gagccattcc tgtggaattt 2340
tcaagcaaca ctgtcaagtt gacgtcgggt catttgaagt gtagagtgaa gatggaaaaa 2400
ttgcagttga agggaacaac ctatggcgtc tgttcaaagg ctttcaagtt tcttgggact 2460
cccgcagaca caggtcacgg cactgtggtg ttggaattgc agtacactgg cacggatgga 2520
ccttgcaaag ttcctatctc gtcagtggct tcattgaacg acctaacgcc agtgggcaga 2580
ttggtcactg tcaacccttt tgtttcagtg gccacggcca acgctaaggt cctgattgaa 2640
ttggaaccac cctttggaga ctcatacata gtggtgggca gaggagaaca acagatcaat 2700
caccattggc acaagtctgg aagcagcatt ggcaaa 2736

<210> 762
<211> 912
<212> PRT
<213> Artificial Sequence

<220>
<223> STF2.E

<400> 762

```
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
 1               5                  10                  15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
            20                  25                  30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
        35                  40                  45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
    50                  55                  60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65                  70                  75                  80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
                85                  90                  95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
            100                 105                 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
        115                 120                 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
    130                 135                 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145                 150                 155                 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val Gln
                165                 170                 175
Lys Ala Tyr Asp Val Lys Asp Thr Ala Val Thr Thr Lys Ala Tyr Ala
```

```
                180                 185                 190
Asn Asn Gly Thr Thr Leu Asp Val Ser Gly Leu Asp Asp Ala Ala Ile
        195                 200                 205
Lys Ala Ala Thr Gly Gly Thr Asn Gly Thr Ala Ser Val Thr Gly Gly
    210                 215                 220
Ala Val Lys Phe Asp Ala Asp Asn Asn Lys Tyr Phe Val Thr Ile Gly
225                 230                 235                 240
Gly Phe Thr Gly Ala Asp Ala Ala Lys Asn Gly Asp Tyr Glu Val Asn
                245                 250                 255
Val Ala Thr Asp Gly Thr Val Thr Leu Ala Ala Gly Ala Thr Lys Thr
            260                 265                 270
Thr Met Pro Ala Gly Ala Thr Thr Lys Thr Glu Val Gln Glu Leu Lys
        275                 280                 285
Asp Thr Pro Ala Val Val Ser Ala Asp Ala Lys Asn Ala Leu Ile Ala
    290                 295                 300
Gly Gly Val Asp Ala Thr Asp Ala Asn Gly Ala Glu Leu Val Lys Met
305                 310                 315                 320
Ser Tyr Thr Asp Lys Asn Gly Lys Thr Ile Glu Gly Gly Tyr Ala Leu
                325                 330                 335
Lys Ala Gly Asp Lys Tyr Tyr Ala Ala Asp Tyr Asp Glu Ala Thr Gly
            340                 345                 350
Ala Ile Lys Ala Lys Thr Thr Ser Tyr Thr Ala Ala Asp Gly Thr Thr
        355                 360                 365
Lys Thr Ala Ala Asn Gln Leu Gly Gly Val Asp Gly Lys Thr Glu Val
    370                 375                 380
Val Thr Ile Asp Gly Lys Thr Tyr Asn Ala Ser Lys Ala Ala Gly His
385                 390                 395                 400
Asp Phe Lys Ala Gln Pro Glu Leu Ala Glu Ala Ala Ala Lys Thr Thr
                405                 410                 415
Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala
            420                 425                 430
Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile
        435                 440                 445
Thr Asn Leu Gly Asn Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg
    450                 455                 460
Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala
465                 470                 475                 480
Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln
                485                 490                 495
Val Pro Gln Asn Val Leu Ser Leu Leu Arg Phe Asn Cys Leu Gly Met
            500                 505                 510
Ser Asn Arg Asp Phe Leu Glu Gly Val Ser Gly Ala Thr Trp Val Asp
        515                 520                 525
Leu Val Leu Glu Gly Asp Ser Cys Val Thr Ile Met Ser Lys Asp Lys
    530                 535                 540
Pro Thr Ile Asp Val Lys Met Met Asn Met Glu Ala Ala Asn Leu Ala
545                 550                 555                 560
Glu Val Arg Ser Tyr Cys Tyr Leu Ala Thr Val Ser Asp Leu Ser Thr
                565                 570                 575
Lys Ala Ala Cys Pro Thr Met Gly Glu Ala His Asn Asp Lys Arg Ala
            580                 585                 590
Asp Pro Ala Phe Val Cys Arg Gln Gly Val Val Asp Arg Gly Trp Gly
        595                 600                 605
Asn Gly Cys Gly Leu Phe Gly Lys Gly Ser Ile Asp Thr Cys Ala Lys
    610                 615                 620
Phe Ala Cys Ser Thr Lys Ala Ile Gly Arg Thr Ile Leu Lys Glu Asn
625                 630                 635                 640
Ile Lys Tyr Glu Val Ala Ile Phe Val His Gly Pro Thr Thr Val Glu
                645                 650                 655
```

```
Ser His Gly Asn Tyr Ser Thr Gln Val Gly Ala Thr Gln Ala Gly Arg
            660                 665                 670
Phe Ser Ile Thr Pro Ala Ala Pro Ser Tyr Thr Leu Lys Leu Gly Glu
        675                 680                 685
Tyr Gly Glu Val Thr Val Asp Cys Glu Pro Arg Ser Gly Ile Asp Thr
    690                 695                 700
Asn Ala Tyr Tyr Val Met Thr Val Gly Thr Lys Thr Phe Leu Val His
705                 710                 715                 720
Arg Glu Trp Phe Met Asp Leu Asn Leu Pro Trp Ser Ser Ala Gly Ser
                725                 730                 735
Thr Val Trp Arg Asn Arg Glu Thr Leu Met Glu Phe Glu Glu Pro His
            740                 745                 750
Ala Thr Lys Gln Ser Val Ile Ala Leu Gly Ser Gln Glu Gly Ala Leu
        755                 760                 765
His Gln Ala Leu Ala Gly Ala Ile Pro Val Glu Phe Ser Ser Asn Thr
    770                 775                 780
Val Lys Leu Thr Ser Gly His Leu Lys Cys Arg Val Lys Met Glu Lys
785                 790                 795                 800
Leu Gln Leu Lys Gly Thr Thr Tyr Gly Val Cys Ser Lys Ala Phe Lys
                805                 810                 815
Phe Leu Gly Thr Pro Ala Asp Thr Gly His Gly Thr Val Val Leu Glu
            820                 825                 830
Leu Gln Tyr Thr Gly Thr Asp Gly Pro Cys Lys Val Pro Ile Ser Ser
        835                 840                 845
Val Ala Ser Leu Asn Asp Leu Thr Pro Val Gly Arg Leu Val Thr Val
    850                 855                 860
Asn Pro Phe Val Ser Val Ala Thr Ala Asn Ala Lys Val Leu Ile Glu
865                 870                 875                 880
Leu Glu Pro Pro Phe Gly Asp Ser Tyr Ile Val Val Gly Arg Gly Glu
                885                 890                 895
Gln Gln Ile Asn His His Trp His Lys Ser Gly Ser Ser Ile Gly Lys
            900                 905                 910
```

<210> 763
<211> 495
<212> PRT
<213> Artificial Sequence

<220>
<223> Dengue 1

<400> 763

```
Met Arg Cys Val Gly Ile Gly Asn Arg Asp Phe Val Glu Gly Leu Ser
 1               5                  10                  15
Gly Ala Thr Trp Val Asp Val Val Leu Glu His Gly Ser Cys Val Thr
            20                  25                  30
Thr Met Ala Lys Asn Lys Pro Thr Leu Asp Ile Glu Leu Leu Lys Thr
        35                  40                  45
Glu Val Thr Asn Pro Ala Val Leu Arg Lys Leu Cys Ile Glu Ala Lys
    50                  55                  60
Ile Ser Asn Thr Thr Thr Asp Ser Arg Cys Pro Thr Gln Gly Glu Ala
65                  70                  75                  80
Thr Leu Val Glu Glu Gln Asp Ala Asn Phe Val Cys Arg Arg Thr Phe
                85                  90                  95
Val Asp Arg Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly Ser
            100                 105                 110
Leu Ile Thr Cys Ala Lys Phe Lys Cys Val Thr Lys Leu Glu Gly Lys
```

```
        115                 120                 125
Ile Ala Gln Tyr Glu Asn Leu Lys Tyr Ser Val Ile Val Thr Val His
    130                 135                 140
Thr Gly Asp Gln His Gln Val Gly Asn Glu Thr Thr Glu His Gly Thr
145                 150                 155                 160
Thr Ala Thr Ile Thr Pro Gln Ala Pro Thr Ser Glu Ile Gln Leu Thr
                165                 170                 175
Asp Tyr Gly Thr Leu Thr Leu Asp Cys Ser Pro Arg Thr Gly Leu Asp
            180                 185                 190
Phe Asn Glu Met Val Leu Leu Thr Met Lys Lys Lys Ser Trp Leu Val
        195                 200                 205
His Lys Gln Trp Phe Leu Asp Leu Pro Leu Pro Trp Thr Ser Gly Ala
    210                 215                 220
Leu Thr Ser Gln Glu Thr Trp Asn Arg Gln Asp Leu Leu Val Thr Phe
225                 230                 235                 240
Lys Thr Ala His Ala Lys Lys Gln Glu Val Val Val Leu Gly Ser Gln
                245                 250                 255
Glu Gly Ala Met His Thr Ala Leu Thr Gly Ala Thr Glu Ile Gln Thr
            260                 265                 270
Ser Gly Thr Thr Thr Ile Phe Ala Gly His Leu Lys Cys Arg Leu Lys
        275                 280                 285
Met Asp Lys Leu Thr Leu Lys Gly Met Ser Tyr Val Met Cys Thr Gly
    290                 295                 300
Ser Phe Lys Leu Glu Lys Glu Val Ala Glu Thr Gln His Gly Thr Val
305                 310                 315                 320
Leu Val Gln Val Lys Tyr Glu Gly Thr Asp Ala Pro Cys Lys Ile Pro
                325                 330                 335
Phe Ser Thr Gln Asp Glu Lys Gly Ala Thr Gln Asn Gly Arg Leu Ile
            340                 345                 350
Thr Ala Asn Pro Ile Val Thr Asp Lys Glu Lys Pro Val Asn Ile Glu
        355                 360                 365
Ala Glu Pro Pro Phe Gly Glu Ser Tyr Ile Val Val Gly Ala Gly Glu
    370                 375                 380
Lys Ala Leu Lys Leu Ser Trp Phe Lys Lys Gly Ser Ser Ile Gly Lys
385                 390                 395                 400
Met Phe Glu Ala Thr Ala Arg Gly Ala Arg Arg Met Ala Ile Leu Gly
                405                 410                 415
Asp Thr Ala Trp Asp Phe Gly Ser Ile Gly Gly Val Phe Thr Ser Met
            420                 425                 430
Gly Lys Leu Val His Gln Val Phe Gly Thr Ala Tyr Gly Val Leu Phe
        435                 440                 445
Ser Gly Val Ser Trp Thr Met Lys Ile Gly Ile Gly Ile Leu Leu Thr
    450                 455                 460
Trp Leu Gly Leu Asn Ser Arg Asn Thr Ser Leu Ser Val Met Cys Ile
465                 470                 475                 480
Ala Val Gly Met Val Thr Leu Tyr Leu Gly Val Met Val Gln Ala
                485                 490                 495
```

<210> 764
<211> 1485
<212> DNA
<213> Artificial Sequence

<220>
<223> Dengue 1

<400> 764

```
atgcgatgcg tgggaatagg caacagagac ttcgtggaag gactgtcagg agcaacatgg 60
gtggatgtgg tactggagca tggaagttgc gtcaccacca tggcaaaaaa caaaccaaca 120
ctggacattg aactcttgaa gacggaggtc acaaaccctg cagttctgcg taaattgtgc 180
attgaagcta aaatatcaaa caccaccacc gattcgagat gtccaacaca aggagaagcc 240
acactggtgg aagaacaaga cgcgaacttt gtgtgccgac gaacgttcgt ggacagaggc 300
tggggcaatg gctgtgggct attcggaaaa ggtagtctaa taacgtgtgc caagtttaag 360
tgtgtgacaa aactagaagg aaagatagct caatatgaaa acctaaaata ttcagtgata 420
gtcaccgtcc acactggaga tcagcaccag gtgggaaatg agactacaga acatggaaca 480
actgcaacca taacacctca agctcctacg tcggaaatac agctgaccga ctacggaacc 540
cttacattag attgttcacc taggacaggg ctagatttta acgagatggt gttgctgaca 600
atgaaaaaga aatcatggct tgtccacaaa cagtggtttc tagacttacc actgccttgg 660
acctctgggg ctttaacatc ccaagagact tggaacagac aagatttact ggtcacattt 720
aagacagctc atgcaaagaa gcaggaagta gtcgtactag gatcacaaga aggagcaatg 780
cacactgcgc tgactggagc gacagaaatc caaacgtcag gaacgacaac aattttcgca 840
ggacacctaa aatgcagact aaaaatggac aaactaactt taaaagggat gtcatatgtg 900
atgtgcacag gctcattcaa gttagagaaa gaagtggctg agacccagca tggaactgtt 960
ctggtgcagg ttaaatatga aggaacagac gcaccatgca agattccctt ttcgacccaa 1020
gatgagaaag gagcaaccca gaatgggaga ttaataacag ccaaccccat agtcactgac 1080
aaagaaaaac cagtcaatat tgaggcagaa ccacccttg gtgagagcta catcgtggta 1140
ggagcaggtg aaaaagcttt gaaactaagc tggttcaaga aaggaagcag catagggaaa 1200
atgtttgaag caactgcccg aggagcacga aggatggcca ttctgggaga caccgcatgg 1260
gacttcggtt ctataggagg agtgttcacg tctatgggaa aactggtaca ccaggttttt 1320
ggaactgcat atggagtttt gtttagcgga gtttcttgga ccatgaaaat aggaataggg 1380
attctgctga catggctagg attaaattca aggaacacgt ccctttcggt gatgtgcatc 1440
gcagttggca tggtcacact gtacctagga gtcatggttc aggca                 1485

<210> 765
<211> 495
<212> PRT
<213> Artificial Sequence

<220>
<223> Dengue 2


<400> 765
Met Arg Cys Ile Gly Met Ser Asn Arg Asp Phe Val Glu Gly Val Ser
 1               5                  10                  15
Gly Gly Ser Trp Val Asp Ile Val Leu Glu His Gly Ser Cys Val Thr
            20                  25                  30
Thr Met Ala Lys Asn Lys Pro Thr Leu Asp Phe Glu Leu Ile Lys Thr
        35                  40                  45
Glu Ala Lys Gln Pro Ala Thr Leu Arg Lys Tyr Cys Ile Glu Ala Lys
    50                  55                  60
Leu Thr Asn Thr Thr Thr Glu Ser Arg Cys Pro Thr Gln Gly Glu Pro
65                  70                  75                  80
Ser Leu Asn Glu Glu Gln Asp Lys Arg Phe Val Cys Lys His Ser Met
                85                  90                  95
Val Asp Arg Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly Gly
            100                 105                 110
Ile Val Thr Cys Ala Met Phe Arg Cys Lys Lys Asn Met Glu Gly Lys
        115                 120                 125
Val Val Gln Pro Glu Asn Leu Glu Tyr Thr Ile Val Ile Thr Pro His
    130                 135                 140
Ser Gly Glu Glu His Ala Val Gly Asn Asp Thr Gly Lys His Gly Lys
145                 150                 155                 160
Glu Ile Lys Ile Thr Pro Gln Ser Ser Thr Thr Glu Ala Glu Leu Thr
                165                 170                 175
Gly Tyr Gly Thr Val Thr Met Glu Cys Ser Pro Arg Thr Gly Leu Asp
```

```
            180                 185                 190
Phe Asn Glu Met Val Leu Leu Gln Met Glu Asn Lys Ala Trp Leu Val
        195                 200                 205
His Arg Gln Trp Phe Leu Asp Leu Pro Leu Pro Trp Leu Pro Gly Ala
    210                 215                 220
Asp Thr Gln Gly Ser Asn Trp Ile Gln Lys Glu Thr Leu Val Thr Phe
225                 230                 235                 240
Lys Asn Pro His Ala Lys Lys Gln Asp Val Val Val Leu Gly Ser Gln
                245                 250                 255
Glu Gly Ala Met His Thr Ala Leu Thr Gly Ala Thr Glu Ile Gln Met
            260                 265                 270
Ser Ser Gly Asn Leu Leu Phe Thr Gly His Leu Lys Cys Arg Leu Arg
        275                 280                 285
Met Asp Lys Leu Gln Leu Lys Gly Met Ser Tyr Ser Met Cys Thr Gly
    290                 295                 300
Lys Phe Lys Val Val Lys Glu Ile Ala Glu Thr Gln His Gly Thr Ile
305                 310                 315                 320
Val Ile Arg Val Gln Tyr Glu Gly Asp Gly Ser Pro Cys Lys Ile Pro
                325                 330                 335
Phe Glu Ile Met Asp Leu Glu Lys Arg His Val Leu Gly Arg Leu Ile
            340                 345                 350
Thr Val Asn Pro Ile Val Thr Glu Lys Asp Ser Pro Val Asn Ile Glu
        355                 360                 365
Ala Glu Pro Pro Phe Gly Asp Ser Tyr Ile Ile Ile Gly Val Glu Pro
    370                 375                 380
Gly Gln Leu Lys Leu Asn Trp Phe Lys Lys Gly Ser Ser Ile Gly Gln
385                 390                 395                 400
Met Phe Glu Thr Thr Met Arg Gly Ala Lys Arg Met Ala Ile Leu Gly
                405                 410                 415
Asp Thr Ala Trp Asp Phe Gly Ser Leu Gly Gly Val Phe Thr Ser Ile
            420                 425                 430
Gly Lys Ala Leu His Gln Val Phe Gly Ala Ile Tyr Gly Ala Ala Phe
        435                 440                 445
Ser Gly Val Ser Trp Thr Met Lys Ile Leu Ile Gly Val Ile Ile Thr
    450                 455                 460
Trp Ile Gly Met Asn Ser Arg Ser Thr Ser Leu Ser Val Thr Leu Val
465                 470                 475                 480
Leu Val Gly Ile Val Thr Leu Tyr Leu Gly Val Met Val Gln Ala
                485                 490                 495
```

<210> 766
<211> 1485
<212> DNA
<213> Artificial Sequence

<220>
<223> Dengue 2

<400> 766
atgcgttgca taggaatgtc aaatagagac tttgtggaag gggtttcagg aggaagctgg 60
gttgacatag tcttagaaca tggaagctgt gtgacgacga tggcaaaaaa caaaccaaca 120
ttggattttg aactgataaa aacagaagcc aaacagcctg ccaccctaag gaagtactgt 180
atagaggcaa agctaaccaa cacaacaaca gaatctcgct gcccaacaca aggggaaccc 240
agcctaaatg aagagcagga caaaaggttc gtctgcaaac actccatggt agacagagga 300
tggggaaatg gatgtggact atttggaaag ggaggcattg tgacctgtgc tatgttcaga 360
tgcaaaaaga acatggaagg aaaagttgtg caaccagaaa acttggaata caccattgtg 420
ataacacctc actcagggga agagcatgca gtcggaaatg acacaggaaa acatggcaag 480

gaaatcaaaa taacaccaca gagttccacc acagaagcag aattgacagg ttatggcact 540
gtcacaatgg agtgctctcc aagaacgggc ctcgacttca atgagatggt gttgctgcag 600
atggaaaata aagcttggct ggtgcacagg caatggttcc tagacctgcc gttaccatgg 660
ttgcccggag cggacacaca agggtcaaat tggatacaga aagagacatt ggtcactttc 720
aaaaatcccc atgcgaagaa acaggatgtt gttgttttag gatcccaaga aggggccatg 780
cacacagcac ttacaggggc cacagaaatc caaatgtcat caggaaactt actcttcaca 840
ggacatctca agtgcaggct gagaatggac aagctacagc tcaaaggaat gtcatactct 900
atgtgcacag gaaagtttaa agttgtgaag gaaatagcag aaacacaaca tggaacaata 960
gttatcagag tgcaatatga aggggacggc tctccatgca agatcccttt tgagataatg 1020
gatttggaaa aaagacatgt cttaggtcgc ctgattacag tcaacccaat tgtgacagaa 1080
aaagatagcc cagtcaacat agaagcagaa cctccattcg gagacagcta catcatcata 1140
ggagtagagc cgggacaact gaagctcaac tggtttaaga aaggaagttc tatcggccaa 1200
atgtttgaga caacaatgag gggggcgaag agaatggcca ttttaggtga cacagcctgg 1260
gattttggat ccttgggagg agtgtttaca tctataggaa aggctctcca ccaagtcttt 1320
ggagcaatct atggagctgc cttcagtggg gtttcatgga ctatgaaaat cctcatagga 1380
gtcattatca catggatagg aatgaattca cgcagcacct cactgtctgt gacactagta 1440
ttggtgggaa ttgtgacact gtatttggga gtcatggtgc aggcc 1485

<210> 767
<211> 493
<212> PRT
<213> Artificial Sequence

<220>
<223> Dengue 3

<400> 767

Met	Arg	Cys	Val	Gly	Val	Gly	Asn	Arg	Asp	Phe	Val	Glu	Gly	Leu	Ser
1				5					10					15	
Gly	Ala	Thr	Trp	Val	Asp	Val	Val	Leu	Glu	His	Gly	Gly	Cys	Val	Thr
			20					25					30		
Thr	Met	Ala	Lys	Asn	Lys	Pro	Thr	Leu	Asp	Ile	Glu	Leu	Gln	Lys	Thr
		35					40					45			
Glu	Ala	Thr	Gln	Leu	Ala	Thr	Leu	Arg	Lys	Leu	Cys	Ile	Glu	Gly	Lys
	50					55					60				
Ile	Thr	Asn	Ile	Thr	Thr	Asp	Ser	Arg	Cys	Pro	Thr	Gln	Gly	Glu	Ala
65					70					75					80
Ile	Leu	Pro	Glu	Glu	Gln	Asp	Gln	Asn	Tyr	Val	Cys	Lys	His	Thr	Tyr
				85					90					95	
Val	Asp	Arg	Gly	Trp	Gly	Asn	Gly	Cys	Gly	Leu	Phe	Gly	Lys	Gly	Ser
			100					105					110		
Leu	Val	Thr	Cys	Ala	Lys	Phe	Gln	Cys	Leu	Glu	Pro	Ile	Glu	Gly	Lys
		115					120					125			
Val	Val	Gln	His	Glu	Asn	Leu	Lys	Tyr	Thr	Val	Ile	Ile	Thr	Val	His
	130					135					140				
Thr	Gly	Asp	Gln	His	Gln	Val	Gly	Asn	Asp	Thr	Gln	Gly	Val	Thr	Val
145					150					155					160
Glu	Ile	Thr	Pro	Gln	Ala	Ser	Thr	Val	Glu	Ala	Ile	Leu	Pro	Glu	Tyr
				165					170					175	
Gly	Thr	Leu	Gly	Leu	Glu	Cys	Ser	Pro	Arg	Thr	Gly	Leu	Asp	Phe	Asn
			180					185					190		
Glu	Met	Ile	Leu	Leu	Thr	Met	Lys	Asn	Lys	Ala	Trp	Met	Val	His	Arg
		195					200					205			
Gln	Trp	Phe	Phe	Asp	Leu	Pro	Leu	Pro	Trp	Thr	Ser	Gly	Ala	Thr	Thr
	210					215					220				
Glu	Thr	Pro	Thr	Trp	Asn	Arg	Lys	Glu	Leu	Leu	Val	Thr	Phe	Lys	Asn
225					230					235					240
Ala	His	Ala	Lys	Lys	Gln	Glu	Val	Val	Val	Leu	Gly	Ser	Gln	Glu	Gly

```
                245                 250                 255
Ala Met His Thr Ala Leu Thr Gly Ala Thr Glu Ile Gln Asn Ser Gly
            260                 265                 270
Gly Thr Ser Ile Phe Ala Gly His Leu Lys Cys Arg Leu Lys Met Asp
        275                 280                 285
Lys Leu Glu Leu Lys Gly Met Ser Tyr Ala Met Cys Leu Asn Thr Phe
    290                 295                 300
Val Leu Lys Lys Glu Val Ser Glu Thr Gln His Gly Thr Ile Leu Ile
305                 310                 315                 320
Lys Val Glu Tyr Lys Gly Glu Asp Ala Pro Cys Lys Ile Pro Phe Ser
                325                 330                 335
Thr Glu Asp Gly Gln Gly Lys Ala His Asn Gly Arg Leu Ile Thr Ala
            340                 345                 350
Asn Pro Val Val Thr Lys Lys Glu Glu Pro Val Asn Ile Glu Ala Glu
        355                 360                 365
Pro Pro Phe Gly Glu Ser Asn Ile Val Ile Gly Ile Gly Asp Lys Ala
    370                 375                 380
Leu Lys Ile Asn Trp Tyr Lys Lys Gly Ser Ser Ile Gly Lys Met Phe
385                 390                 395                 400
Glu Ala Thr Ala Arg Gly Ala Arg Arg Met Ala Ile Leu Gly Asp Thr
                405                 410                 415
Ala Trp Asp Phe Gly Ser Val Gly Gly Val Leu Asn Ser Leu Gly Lys
            420                 425                 430
Met Val His Gln Ile Phe Gly Ser Ala Tyr Thr Ala Leu Phe Ser Gly
        435                 440                 445
Val Ser Trp Ile Met Lys Ile Gly Ile Gly Val Leu Leu Thr Trp Ile
    450                 455                 460
Gly Leu Asn Ser Lys Asn Thr Ser Met Ser Phe Ser Cys Ile Ala Ile
465                 470                 475                 480
Gly Ile Ile Thr Leu Tyr Leu Gly Ala Val Val Gln Ala
                485                 490
```

<210> 768
<211> 1479
<212> DNA
<213> Artificial Sequence

<220>
<223> Dengue 3

<400> 768

```
atgagatgcg tgggagtagg aaacagagat tttgtggaag gtctgtcggg agctacgtgg 60
gttgatgtgg tgctcgagca cggagggtgt gtgaccacca tggctaagaa caagcctacg 120
ctggacatag agctccagaa gaccgaggcc acccaactgg cgaccctaag aaagttatgc 180
attgagggaa aaattaccaa cataacaact gactcaaggt gtcctaccca gggggaagcg 240
attttacctg aggagcagga ccagaactac gtatgtaagc acacatacgt ggatagaggt 300
tggggaaacg gttgtggttt gtttggaaaa ggaagcttgg tgacatgcgc gaaatttcaa 360
tgtctagaac caatagaggg aaaagtggtg caacatgaga acctcaaata cactgtcatc 420
atcacagtgc acacaggaga ccaacaccag gtgggaaatg acacgcaggg agtcacggtt 480
gagataacac cccaggcatc aaccgttgaa gctatcttgc ctgaatatgg aacccttggg 540
ctagaatgtt caccacggac aggtttggat ttcaacgaaa tgattttatt gacaatgaag 600
aacaaagcat ggatggtaca taggcaatgg ttctttgacc tacccctacc atggacatca 660
ggagctacaa cagagacacc aacttggaac aggaaagagc tccttgtgac attcaagaat 720
gcacatgcaa aaaagcaaga agtagttgtc cttggatcgc aagagggagc aatgcacaca 780
gcgctgacag gagctacaga gattcaaaac tcaggaggta caagcatttt tgcggggcac 840
ttgaaatgta gacttaagat ggacaaattg gaactcaagg ggatgagcta tgcaatgtgc 900
ttgaatacct ttgtgttgaa gaaagaagtc tctgaaacgc agcatgggac aatactcatc 960
```

```
aaggttgagt acaaagggga agatgcacct tgcaagattc ctttctccac agaggatgga 1020
caagggaaag ctcacaatgg tagactgatc acagccaacc cagtggtgac caagaaggag 1080
gagcctgtca acattgaggc tgaacctcct tttggggaaa gtaacatagt gattgggatt 1140
ggagacaaag ccttgaaaat taactggtac aagaagggaa gctcgattgg aaagatgttc 1200
gaggctactg ccagaggtgc aaggcgcatg gccatcttgg gagacacagc ctgggatttt 1260
ggttcagtgg gtggtgttct gaattcatta gggaaaatgg tacaccaaat attcggaagt 1320
gcttacacag ccctgtttag tggagtctca tggataatga aaattggaat aggtgtcctc 1380
ttaacctgga tagggttgaa ttcaaaaaac acttccatgt cattttcatg cattgcgata 1440
ggaattatta cactctatct gggagccgtg gtacaagct                        1479

<210> 769
<211> 491
<212> PRT
<213> Artificial Sequence

<220>
<223> Dengue 4


<400> 769
Met Arg Cys Val Gly Val Gly Asn Arg Asp Phe Val Glu Gly Val Ser
 1               5                  10                  15
Gly Gly Ala Trp Val Asp Leu Val Leu Glu His Gly Gly Cys Val Thr
            20                  25                  30
Thr Met Ala Gln Gly Lys Pro Thr Leu Asp Phe Glu Leu Thr Lys Thr
        35                  40                  45
Thr Ala Lys Glu Val Ala Leu Leu Arg Thr Tyr Cys Ile Glu Ala Ser
    50                  55                  60
Ile Ser Asn Ile Thr Thr Ala Thr Arg Cys Pro Thr Gln Gly Glu Pro
65                  70                  75                  80
Tyr Leu Lys Glu Glu Gln Asp Gln Gln Tyr Ile Cys Arg Arg Asp Val
                85                  90                  95
Val Asp Arg Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly Gly
            100                 105                 110
Val Val Thr Cys Ala Lys Phe Ser Cys Ser Gly Lys Ile Thr Gly Asn
        115                 120                 125
Leu Val Arg Ile Glu Asn Leu Glu Tyr Thr Val Val Val Thr Val His
    130                 135                 140
Asn Gly Asp Thr His Ala Val Gly Asn Asp Thr Ser Asn His Gly Val
145                 150                 155                 160
Thr Ala Met Ile Thr Pro Arg Ser Pro Ser Val Glu Val Lys Leu Pro
                165                 170                 175
Asp Tyr Gly Glu Leu Thr Leu Asp Cys Glu Pro Arg Ser Gly Ile Asp
            180                 185                 190
Phe Asn Glu Met Ile Leu Met Lys Met Lys Lys Lys Thr Trp Leu Val
        195                 200                 205
His Lys Gln Trp Phe Leu Asp Leu Pro Leu Pro Trp Thr Ala Gly Ala
    210                 215                 220
Asp Thr Ser Glu Val His Trp Asn Tyr Lys Glu Arg Met Val Thr Phe
225                 230                 235                 240
Lys Val Pro His Ala Lys Arg Gln Asp Val Thr Val Leu Gly Ser Gln
                245                 250                 255
Glu Gly Ala Met His Ser Ala Leu Ala Gly Ala Thr Glu Val Asp Ser
            260                 265                 270
Gly Asp Gly Asn His Met Phe Ala Gly His Leu Lys Cys Glu Val Arg
        275                 280                 285
Met Glu Lys Leu Arg Ile Lys Gly Met Ser Tyr Thr Met Cys Ser Gly
    290                 295                 300
Lys Phe Ser Ile Asp Lys Glu Met Ala Glu Thr Gln His Gly Thr Thr
```

```
305                 310                 315                 320
Val Val Lys Val Lys Tyr Glu Gly Ala Gly Ala Pro Cys Lys Val Pro
                325                 330                 335
Ile Glu Ile Arg Asp Val Asn Lys Glu Lys Val Val Gly Arg Ile Ile
            340                 345                 350
Ser Ser Thr Pro Leu Ala Glu Asn Thr Asn Ser Val Thr Asn Ile Glu
        355                 360                 365
Leu Glu Pro Pro Phe Gly Asp Ser Tyr Ile Val Ile Gly Val Gly Asn
    370                 375                 380
Ser Ala Leu Thr Leu His Trp Phe Arg Lys Gly Ser Ser Ile Gly Lys
385                 390                 395                 400
Met Phe Glu Ser Thr Tyr Arg Gly Ala Lys Arg Met Ala Ile Leu Gly
                405                 410                 415
Glu Thr Ala Trp Asp Phe Gly Ser Val Gly Gly Leu Phe Thr Ser Leu
            420                 425                 430
Gly Lys Ala Val His Gln Val Phe Gly Ser Val Tyr Thr Thr Met Phe
        435                 440                 445
Gly Gly Val Ser Trp Met Ile Arg Ile Leu Ile Gly Phe Leu Val Leu
    450                 455                 460
Trp Ile Gly Thr Asn Ser Arg Asn Thr Ser Met Ala Met Thr Cys Ile
465                 470                 475                 480
Ala Val Gly Gly Ile Thr Leu Phe Leu Gly Phe
                485                 490
```

<210> 770
<211> 1473
<212> DNA
<213> Artificial Sequence

<220>
<223> Dengue 4

<400> 770

```
atgcgatgcg taggagtagg aaacagagac tttgtggaag gagtctcagg tggagcatgg 60
gtcgacctgg tgctagaaca tggaggatgc gtcacaacca tggcccaggg aaaaccaacc 120
ttggattttg aactgaccaa gacaacagcc aaggaagtgg ctctgttaag aacctattgc 180
attgaagcct caatatcaaa cataactacg gcaacaagat gtccaacgca aggagagcct 240
tatctgaaag aggaacagga ccaacagtac atttgccgga gagatgtggt agacagaggg 300
tggggcaatg gctgtggctt gtttggaaaa ggaggagttg tgacatgtgc gaagttttca 360
tgttcgggga agataacagg caatttggtc cgaattgaga accttgaata cacagtggtg 420
gtaacagtcc acaatggaga cacccatgca gtaggaaatg acacatccaa tcatggagtt 480
acagccatga taacccccag gtcaccatcg gtggaagtca aattgccgga ctatggagaa 540
ctaacactcg attgtgaacc caggtctgga attgacttta atgagatgat tctgatgaaa 600
atgaaaaaga aaacatggct cgtgcataag caatggtttt tggatctgcc tcttccatgg 660
acagcaggag cagacacatc agaggttcac tggaattaca aagagagaat ggtgacattt 720
aaggttcctc atgccaagag acaggatgtg acagtgctgg gatctcagga aggagccatg 780
cattctgccc tcgctggagc cacagaagtg gactccggtg atggaaatca catgtttgca 840
ggacatctca agtgcgaagt ccgtatggag aaattgagaa tcaagggaat gtcatacacg 900
atgtgttcag gaaagttttc aattgacaaa gagatggcag aaacacagca tgggacaaca 960
gtggtgaaag tcaagtatga aggtgctgga gctccgtgta aagtccccat agagataaga 1020
gatgtaaaca aggaaaaagt ggttgggcgt atcatctcat ccaccccttt ggctgagaat 1080
accaacagtg taaccaacat agaattagaa ccccccttg gggacagcta catagtgata 1140
ggtgttggaa acagcgcatt aacactccat tggttcagga aagggagttc cattggcaag 1200
atgtttgagt ccacatacag aggtgcaaaa cgaatggcca ttctaggtga aacagcttgg 1260
gattttggtt ccgttggtgg attgttcaca tcattgggaa aggctgtgca ccaggttttt 1320
ggaagtgtgt atacaaccat gtttggagga gtctcatgga tgattagaat cctaattggg 1380
ttcttagtgt tgtggattgg cacgaactca aggaacactt caatggctat gacgtgcata 1440
```

gctgttggag gaatcactct gtttctgggc ttc 1473

<210> 771
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> West Nile virus

<400> 771
Leu Thr Ser Gly His Leu Lys Cys Arg Val Lys Met Glu Lys Leu Gln
1 5 10 15
Leu Lys Gly Thr
20

<210> 772
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> West Nile virus

<400> 772
Leu Thr Ser Gly His Leu Lys Cys Arg Val Lys Met Glu Lys Leu Gln
1 5 10 15
Leu Lys Gly Thr
20

<210> 773
<211> 1217
<212> DNA
<213> Artificial Sequence

<220>
<223> Japanese encephalitis

<400> 773
tttaattgtc tgggaatggg caatcgtgac ttcatagaag gagccagtgg agccgcttgg 60
gtgacttggt gctagaagga gacagctgct tgacaatcat ggcaaacgac aaaccaacat 120
tggacgtccg catgattaac atcgaagcta gccaacttgc tgaggtcaga agttactgct 180
atcatgcttc agtcactgac atctcgacgg tggctcggtg ccccacgact ggagaagccc 240
acaacgagaa gcgagctgat agtagctatg tgtgcaaaca aggcttcact gaccgtgggt 300
ggggcaacgg atgtggattt ttcgggaagg gaagcattga cacatgtgca aaattctcct 360
gcaccagtaa agcgattggg agaacaatcc agccagaaaa catcaaatac aaagttggca 420
tttttgtgca tggaaccacc acttcggaaa accatgggaa ttattcagcg caagttgggg 480
cgtcccaggc ggcaaagttt acagtaacac ccaatgctcc ttcggtagcc ctcaaacttg 540
gtgactacgg agaagtcaca ctggactgtg agccaaggag tggactgaac actgaagcgt 600
tttacgtcat gaccgtgggg tcaaagtcat ttctggtcca tagggagtgg tttcatgacc 660
tcgctctccc ctggacgtcc ccttcgagca cagcgtggag aaacagagaa ctcctcatgg 720
aatttgaagg ggcgcacgcc acaaaacagt ccgttgttgc tcttgggtca caggaaggag 780
gcctccatca tgcgttggca ggagccatcg tggtggagta ctcaagctca gtgatgttaa 840
catcaggcca cctgaaatgt aggctgaaaa tggacaaact ggctctgaaa ggcacaacct 900

atggcatgtg tacagaaaaa ttctcgttcg cgaaaaatcc ggtggacact ggtcacggaa 960
cagttgtcat tgaactctcc tactctggga gtgatggccc ctgcaaaatt ccgattgttt 1020
ccgttgcgag cctcaatgac atgacccccg ttgggcggct ggtgacagtg aacccttcg 1080
tcgcgacttc cagtgccaac tcaaaggtgc tggtcgagat ggaacccccc ttcggagact 1140
cctacatcgt agttggaagg ggagacaagc agatcaacca ccattggcac aaagctggaa 1200
gcacgctggg caaggcc 1217

<210> 774
<211> 406
<212> PRT
<213> Artificial Sequence

<220>
<223> Japanese encephalitis

<400> 774

Phe	Asn	Cys	Leu	Gly	Met	Gly	Asn	Arg	Asp	Phe	Ile	Glu	Gly	Ala	Ser
1				5					10					15	
Gly	Ala	Ala	Trp	Val	Asp	Leu	Val	Leu	Glu	Gly	Asp	Ser	Cys	Leu	Thr
			20					25					30		
Ile	Met	Ala	Asn	Asp	Lys	Pro	Thr	Leu	Asp	Val	Arg	Met	Ile	Asn	Ile
		35					40					45			
Glu	Ala	Ser	Gln	Leu	Ala	Glu	Val	Arg	Ser	Tyr	Cys	Tyr	His	Ala	Ser
	50					55					60				
Val	Thr	Asp	Ile	Ser	Thr	Val	Ala	Arg	Cys	Pro	Thr	Thr	Gly	Glu	Ala
65					70					75					80
His	Asn	Glu	Lys	Arg	Ala	Asp	Ser	Ser	Tyr	Val	Cys	Lys	Gln	Gly	Phe
				85					90					95	
Thr	Asp	Arg	Gly	Trp	Gly	Asn	Gly	Cys	Gly	Phe	Phe	Gly	Lys	Gly	Ser
			100					105					110		
Ile	Asp	Thr	Cys	Ala	Lys	Phe	Ser	Cys	Thr	Ser	Lys	Ala	Ile	Gly	Arg
		115					120					125			
Thr	Ile	Gln	Pro	Glu	Asn	Ile	Lys	Tyr	Lys	Val	Gly	Ile	Phe	Val	His
	130					135					140				
Gly	Thr	Thr	Thr	Ser	Glu	Asn	His	Gly	Asn	Tyr	Ser	Ala	Gln	Val	Gly
145					150					155					160
Ala	Ser	Gln	Ala	Ala	Lys	Phe	Thr	Val	Thr	Pro	Asn	Ala	Pro	Ser	Val
				165					170					175	
Ala	Leu	Lys	Leu	Gly	Asp	Tyr	Gly	Glu	Val	Thr	Leu	Asp	Cys	Glu	Pro
			180					185					190		
Arg	Ser	Gly	Leu	Asn	Thr	Glu	Ala	Phe	Tyr	Val	Met	Thr	Val	Gly	Ser
		195					200					205			
Lys	Ser	Phe	Leu	Val	His	Arg	Glu	Trp	Phe	His	Asp	Leu	Ala	Leu	Pro
	210					215					220				
Trp	Thr	Ser	Pro	Ser	Ser	Thr	Ala	Trp	Arg	Asn	Arg	Glu	Leu	Leu	Met
225					230					235					240
Glu	Phe	Glu	Gly	Ala	His	Ala	Thr	Lys	Gln	Ser	Val	Val	Ala	Leu	Gly
				245					250					255	
Ser	Gln	Glu	Gly	Gly	Leu	His	His	Ala	Leu	Ala	Gly	Ala	Ile	Val	Val
			260					265					270		
Glu	Tyr	Ser	Ser	Ser	Val	Met	Leu	Thr	Ser	Gly	His	Leu	Lys	Cys	Arg
		275					280					285			
Leu	Lys	Met	Asp	Lys	Leu	Ala	Leu	Lys	Gly	Thr	Thr	Tyr	Gly	Met	Cys
	290					295					300				
Thr	Glu	Lys	Phe	Ser	Phe	Ala	Lys	Asn	Pro	Val	Asp	Thr	Gly	His	Gly
305					310					315					320
Thr	Val	Val	Ile	Glu	Leu	Ser	Tyr	Ser	Gly	Ser	Asp	Gly	Pro	Cys	Lys
				325					330					335	

```
Ile Pro Ile Val Ser Val Ala Ser Leu Asn Asp Met Thr Pro Val Gly
            340                 345                 350
Arg Leu Val Thr Val Asn Pro Phe Val Ala Thr Ser Ser Ala Asn Ser
        355                 360                 365
Lys Val Leu Val Glu Met Glu Pro Pro Phe Gly Asp Ser Tyr Ile Val
    370                 375                 380
Val Gly Arg Gly Asp Lys Gln Ile Asn His His Trp His Lys Ala Gly
385                 390                 395                 400
Ser Thr Leu Gly Lys Ala
                405
```

<210> 775
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> West Nile virus

<400> 775
Met Glu Lys Leu Gln
 1   5

<210> 776
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> West Nile virus

<400> 776
Met Asp Lys Leu Ala Leu
 1   5

<210> 777
<211> 496
<212> PRT
<213> Artificial Sequence

<220>
<223> Tick-borne encephalitis

<400> 777

```
Ser Arg Cys Thr His Leu Glu Asn Arg Asp Phe Val Thr Gly Thr Gln
 1               5                  10                  15
Gly Thr Thr Arg Val Thr Leu Val Leu Glu Leu Gly Gly Cys Val Thr
            20                  25                  30
Ile Thr Ala Glu Gly Lys Pro Ser Met Asp Val Trp Leu Asp Ser Ile
        35                  40                  45
Tyr Gln Glu Asn Pro Ala Lys Thr Arg Glu Tyr Cys Leu His Ala Lys
    50                  55                  60
Leu Ser Asp Thr Lys Val Ala Ala Arg Cys Pro Thr Met Gly Pro Ala
```

65 70 75 80
Thr Leu Thr Glu Glu His Gln Ser Gly Thr Val Cys Lys Arg Asp Gln
85 90 95
Ser Asp Arg Gly Trp Gly Asn His Cys Gly Leu Phe Gly Lys Gly Ser
100 105 110
Ile Val Thr Cys Val Lys Val Ala Cys Glu Ala Lys Lys Lys Ala Ile
115 120 125
Gly His Val Tyr Asp Ala Asn Lys Ile Val Tyr Thr Val Lys Val Glu
130 135 140
Pro His Thr Gly Asp Tyr Val Ala Ala Asn Glu Thr His Ser Gly Arg
145 150 155 160
Lys Thr Ala Ser Phe Thr Val Ser Ser Glu Lys Thr Ile Leu Thr Met
165 170 175
Gly Asp Tyr Gly Asp Val Pro Leu Leu Cys Arg Val Ala Ser Gly Val
180 185 190
Asp Leu Ala Gln Thr Val Ile Leu Glu Leu Asp Lys Thr Leu Glu His
195 200 205
Leu Pro Thr Ala Trp Gln Val His Arg Asp Trp Phe Asn Asp Leu Ala
210 215 220
Leu Pro Trp Lys His Glu Gly Ala Gln Gln Trp Asn Asn Ala Glu Arg
225 230 235 240
Leu Val Glu Phe Gly Ala Pro His Ala Val Lys Met Asp Val Tyr Asn
245 250 255
Leu Gly Asp Gln Thr Gly Val Leu Leu Lys Ser Leu Ala Gly Val Pro
260 265 270
Val Ala His Ile Asp Gly Thr Lys Tyr His Leu Lys Ser Gly His Val
275 280 285
Thr Cys Glu Val Gly Leu Glu Lys Leu Lys Met Lys Gly Leu Thr Tyr
290 295 300
Thr Met Cys Asp Lys Thr Lys Phe Ala Trp Lys Arg Thr Pro Thr Asp
305 310 315 320
Ser Gly His Asp Thr Val Val Met Glu Val Thr Phe Ser Gly Thr Lys
325 330 335
Pro Cys Arg Ile Pro Val Arg Ala Val Ala His Gly Ser Pro Asp Val
340 345 350
Asn Val Ala Met Leu Ile Thr Pro Asn Pro Thr Ile Glu Asn Asn Gly
355 360 365
Gly Gly Phe Ile Glu Met Gln Leu Pro Pro Gly Asp Asn Ile Ile Tyr
370 375 380
Val Gly Glu Leu Ser His Gln Trp Phe Gln Lys Gly Ser Ser Ile Gly
385 390 395 400
Arg Val Phe Gln Lys Thr Arg Lys Gly Ile Glu Arg Leu Thr Val Ile
405 410 415
Gly Glu His Ala Trp Asp Phe Gly Ser Thr Gly Gly Phe Leu Thr Ser
420 425 430
Val Gly Lys Ala Leu His Thr Val Leu Gly Gly Ala Phe Asn Ser Ile
435 440 445
Phe Gly Gly Val Gly Phe Leu Pro Lys Leu Leu Leu Gly Val Ala Leu
450 455 460
Ala Trp Leu Gly Leu Asn Met Arg Asn Pro Thr Met Ser Met Ser Phe
465 470 475 480
Leu Leu Ala Gly Gly Leu Val Leu Ala Met Thr Leu Gly Val Gly Ala
485 490 495

<210> 778
<211> 1488
<212> DNA
<213> Artificial Sequence

<220>
<223> Tick-borne encephalitis

<400> 778
tcgcggtgca cacatttgga gaacagggac tttgtcactg gtactcaggg aaccacgaga 60
gtgactctgg tgttggagct gggggggatgt gtcacgatca ctgctgaggg gaagccctca 120
atggatgtgt ggctcgattc catctatcag gagaaccctg ccaagacacg cgagtactgt 180
ctgcatgcca agctgtcgga caccaaagtt gcggccagat gcccaacaat ggggcctgcc 240
actctgactg aggagcatca gagtggtacg gtgtgcaaga gagaccagag tgaccgaggc 300
tggggcaatc actgcggatt gtttggaaag ggcagtattg tgacctgtgt caaggtggct 360
tgtgaggcaa agaagaaggc cattggacat gtgtatgatg ccaacaagat cgtgtacacc 420
gttaaggttg agccacacac gggggactat gttgccgcca atgaaaccca cagtgggagg 480
aagacggcat ccttcacggt ctcctcagaa aagaccatct tgactatggg ggactatgga 540
gatgtgccct tgttgtgcag agtcgccagt ggcgttgact tggctcagac tgtcattctt 600
gagcttgaca agactctgga acaccttcca acagcctggc aggtccatcg tgactggttc 660
aatgatctgg ctctaccgtg gaaacacgaa ggagcgcaac aatggaacaa tgctgagcga 720
ctggttgaat ttggagctcc gcatgccgtt aaaatggacg tgtataacct tggagatcaa 780
actggggtgt tgttgaagtc acttgctggg gttcctgtgg cgcacattga tggaacaaag 840
taccacctaa aaagcggcca tgtaacatgc gaggttggac tagaaaagct caaaatgaag 900
ggtctcacat acacaatgtg tgacaaaacg aagttcgcat ggaagcggac tccaacagac 960
agcggacatg acacagtggt catggaggtc acgttctctg gaacaaaacc ttgcaggatc 1020
ccagtgcggg cagtggcaca cggctctcca gatgtaaatg tggccatgct gataacacca 1080
aaccccacca ttgagaacaa tggaggtggc ttcatagaga tgcagctacc cccaggggat 1140
aacatcatct atgttgggga actaagccat cagtggttcc agaaggggag cagcatcgga 1200
agggtgtttc aaaagaccag gaagggcatc gagagactga cagtgatagg agaacacgcc 1260
tgggacttcg gttccactgg aggtttcttg acttcggtag gcaaagcgct gcacacagtc 1320
ctcggcggag ccttcaacag catctttggg ggagtggggt ttctgcccaa gctcctgttg 1380
ggtgtggcct tagcctggtt gggcctgaac atgaggaacc ccaccatgtc catgagtttc 1440
ctcttggctg ggggactggt cctggctatg acacttggag tgggtgct 1488

<210> 779
<211> 110
<212> PRT
<213> Artificial Sequence

<220>
<223> Domain III of West Nile virus

<400> 779
Leu Lys Gly Thr Thr Tyr Gly Val Cys Ser Lys Ala Phe Lys Phe Leu
1 5 10 15
Gly Thr Pro Ala Asp Thr Gly His Gly Thr Val Val Leu Glu Leu Gln
20 25 30
Tyr Thr Gly Thr Asp Gly Pro Cys Lys Val Pro Ile Ser Ser Val Ala
35 40 45
Ser Leu Asn Asp Leu Thr Pro Val Gly Arg Leu Val Thr Val Asn Pro
50 55 60
Phe Val Ser Val Ala Thr Ala Asn Ala Lys Val Leu Ile Glu Leu Glu
65 70 75 80
Pro Pro Phe Gly Asp Ser Tyr Ile Val Val Gly Arg Gly Glu Gln Gln
85 90 95
Ile Asn His His Trp His Lys Ser Gly Ser Ser Ile Gly Lys
100 105 110

<210> 780

<211> 110
<212> PRT
<213> Artificial Sequence

<220>
<223> Domain III of Japanese encephalitis virus

<400> 780
Lys Gly Thr Thr Tyr Gly Met Cys Thr Glu Lys Phe Ser Phe Ala Lys
1 5 10 15
Asn Pro Val Asp Thr Gly His Gly Thr Val Val Ile Glu Leu Ser Tyr
20 25 30
Ser Gly Ser Asp Gly Pro Cys Lys Ile Pro Ile Val Ser Val Ala Ser
35 40 45
Leu Asn Asp Met Thr Pro Val Gly Arg Leu Val Thr Val Asn Pro Phe
50 55 60
Val Ala Thr Ser Ser Ala Asn Ser Lys Val Leu Val Glu Met Glu Pro
65 70 75 80
Pro Phe Gly Asp Ser Tyr Ile Val Val Gly Arg Gly Asp Lys Gln Ile
85 90 95
Asn His His Trp His Lys Ala Gly Ser Thr Leu Gly Lys Ala
100 105 110

<210> 781
<211> 345
<212> DNA
<213> Artificial Sequence

<220>
<223> West Nile EIII+

<400> 781
atggaaaaat tgcagttgaa gggaacaacc tatggcgtct gttcaaaggc tttcaagttt 60
cttgggactc ccgcagacac aggtcacggc actgtggtgt tggaattgca gtacactggc 120
acggatggac cttgcaaagt tcctatctcg tcagtggctt cattgaacga cctaacgcca 180
gtgggcagat tggtcactgt caacccttttt gtttcagtgg ccacggccaa cgctaaggtc 240
ctgattgaat tggaaccacc ctttggagac tcatacatag tggtgggcag aggagaacaa 300
cagatcaatc accattggca caagtctgga agcagcattg gcaaa 345

<210> 782
<211> 23
<212> PRT
<213> Artificial Sequence

<220>
<223> P. aeruginosa

<400> 782
Met Asn Asn Val Leu Lys Phe Ser Ala Leu Ala Leu Ala Ala Val Leu
1 5 10 15
Ala Thr Gly Cys Ser Ser His
20

<210> 783

<211> 72
<212> DNA
<213> Artificial Sequence

<220>
<223> P. aeruginosa

<400> 783
atgaaagcta ctaaactggt actgggcgcg gtaatcctgg gttctactct gctggcaggt 60
tgctccagca ac 72

<210> 784
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> E. coli

<400> 784
Met Lys Ala Thr Lys Leu Val Leu Gly Ala Val Ile Leu Gly Ser Thr
1 5 10 15
Leu Leu Ala Gly Cys Ser Ser Asn
20

<210> 785
<211> 72
<212> DNA
<213> Artificial Sequence

<220>
<223> E. coli

<400> 785
atgaaagcta ctaaactggt actgggcgcg gtaatcctgg gttctactct gctggcaggt 60
tgctccagca ac 72

<210> 786
<211> 11
<212> PRT
<213> Artificial Sequence

<220>
<223> Linker

<400> 786
His Gly Ala Pro Val Asp Pro Ala Ser Pro Trp
1 5 10

<210> 787
<211> 585
<212> PRT
<213> B/Victoria/2/87 HA

<400> 787
Met Lys Ala Ile Ile Val Leu Leu Met Val Val Thr Ser Asn Ala Asp
1 5 10 15
Arg Ile Cys Thr Gly Ile Thr Ser Ser Asn Ser Pro His Val Val Lys
20 25 30
Thr Ala Thr Gln Gly Glu Val Asn Val Thr Gly Val Ile Pro Leu Thr
35 40 45
Thr Thr Pro Thr Lys Ser His Phe Ala Asn Leu Lys Gly Thr Lys Thr
50 55 60
Arg Gly Lys Leu Cys Pro Lys Cys Leu Asn Cys Thr Asp Leu Asp Val
65 70 75 80
Ala Leu Gly Arg Pro Lys Cys Thr Gly Thr Ile Pro Ser Ala Lys Ala
85 90 95
Ser Ile Leu His Glu Val Lys Pro Val Thr Ser Gly Cys Phe Pro Ile
100 105 110
Met His Asp Arg Thr Lys Ile Arg Gln Leu Pro Asn Leu Leu Arg Gly
115 120 125
Tyr Glu His Ile Arg Leu Ser Thr His Asn Val Ile Asn Ala Glu Thr
130 135 140
Ala Pro Gly Gly Pro Tyr Lys Val Gly Thr Ser Gly Ser Cys Pro Asn
145 150 155 160
Val Thr Asn Gly Asn Gly Phe Phe Ala Thr Met Ala Trp Ala Val Pro
165 170 175
Lys Asn Asp Asn Asn Lys Thr Ala Thr Asn Pro Leu Thr Val Glu Val
180 185 190
Pro Tyr Ile Cys Thr Glu Gly Glu Asp Gln Ile Thr Val Trp Gly Phe
195 200 205
His Ser Asp Asn Glu Ala Gln Met Val Lys Leu Tyr Gly Asp Ser Lys
210 215 220
Pro Gln Lys Phe Thr Ser Ser Ala Asn Gly Val Thr Thr His Tyr Val
225 230 235 240
Ser Gln Ile Gly Gly Phe Pro Asn Gln Ala Glu Asp Gly Gly Leu Pro
245 250 255
Gln Ser Gly Arg Ile Val Val Asp Tyr Met Val Gln Lys Ser Gly Lys
260 265 270
Thr Gly Thr Ile Thr Tyr Gln Arg Gly Ile Leu Leu Pro Gln Lys Val
275 280 285
Trp Cys Ala Ser Gly Arg Ser Lys Val Ile Lys Gly Ser Leu Pro Leu
290 295 300
Ile Gly Glu Ala Asp Cys Leu His Glu Lys Tyr Gly Gly Leu Asn Lys
305 310 315 320
Ser Lys Pro Tyr Tyr Thr Gly Glu His Ala Lys Ala Ile Gly Asn Cys
325 330 335
Pro Ile Trp Val Lys Thr Pro Leu Lys Leu Ala Asn Gly Thr Lys Tyr
340 345 350
Arg Pro Pro Ala Lys Leu Leu Lys Glu Lys Gly Phe Phe Gly Ala Ile
355 360 365
Ala Gly Phe Leu Glu Gly Gly Trp Glu Gly Met Ile Ala Gly Trp His
370 375 380
Gly Tyr Thr Ser His Gly Ala His Gly Val Ala Val Ala Ala Asp Leu
385 390 395 400
Lys Ser Thr Gln Glu Ala Ile Asn Lys Ile Thr Lys Asn Leu Asn Ser
405 410 415
Leu Ser Glu Leu Glu Val Lys Asn Leu Gln Arg Leu Ser Gly Ala Met
420 425 430
Asp Glu Leu His Asn Lys Ile Leu Glu Leu Asp Glu Lys Val Asp Asp
435 440 445
Leu Arg Ala Asp Thr Ile Ser Ser Gln Ile Glu Leu Ala Val Leu Leu

```
    450                 455                 460
Ser Asn Glu Gly Ile Ile Asn Ser Glu Asp Glu His Leu Leu Ala Leu
465                 470                 475                 480
Glu Arg Lys Leu Lys Lys Met Leu Gly Pro Ser Ala Val Glu Ile Gly
                485                 490                 495
Asn Gly Cys Phe Glu Thr Lys His Lys Cys Asn Gln Thr Cys Leu Asp
            500                 505                 510
Arg Ile Ala Ala Gly Thr Phe Asn Ala Gly Glu Phe Ser Leu Pro Thr
        515                 520                 525
Phe Asp Ser Leu Asn Ile Thr Ala Ala Ser Leu Asn Asp Asp Gly Leu
    530                 535                 540
Asp Asn His Thr Ile Leu Leu Tyr Tyr Ser Thr Ala Ala Ser Ser Leu
545                 550                 555                 560
Ala Val Thr Leu Met Ile Ala Ile Phe Ile Val Tyr Met Val Ser Arg
                565                 570                 575
Asp Asn Val Ser Cys Ser Ile Cys Leu
            580                 585
```

<210> 788
<211> 535
<212> PRT
<213> Artificial Sequence

<220>
<223> Cross-reactive mutant (CRM) of diphtheria toxin including, CRM197

<400> 788

```
Gly Ala Asp Asp Val Val Asp Ser Ser Lys Ser Phe Val Met Glu Asn
 1                   5                  10                  15
Phe Ser Ser Tyr His Gly Thr Lys Pro Gly Tyr Val Asp Ser Ile Gln
            20                  25                  30
Lys Gly Ile Gln Lys Pro Lys Ser Gly Thr Gln Gly Asn Tyr Asp Asp
        35                  40                  45
Asp Trp Lys Gly Phe Tyr Ser Thr Asp Asn Lys Tyr Asp Ala Ala Gly
    50                  55                  60
Tyr Ser Val Asp Asn Glu Asn Pro Leu Ser Gly Lys Ala Gly Gly Val
65                  70                  75                  80
Val Lys Val Thr Tyr Pro Gly Leu Thr Lys Val Leu Ala Leu Lys Val
                85                  90                  95
Asp Asn Ala Glu Thr Ile Lys Lys Glu Leu Gly Leu Ser Leu Thr Glu
            100                 105                 110
Pro Leu Met Glu Gln Val Gly Thr Glu Glu Phe Ile Lys Arg Phe Gly
        115                 120                 125
Asp Gly Ala Ser Arg Val Val Leu Ser Leu Pro Phe Ala Glu Gly Ser
    130                 135                 140
Ser Ser Val Glu Tyr Ile Asn Asn Trp Glu Gln Ala Lys Ala Leu Ser
145                 150                 155                 160
Val Glu Leu Glu Ile Asn Phe Glu Thr Arg Gly Lys Arg Gly Gln Asp
                165                 170                 175
Ala Met Tyr Glu Tyr Met Ala Gln Ala Cys Ala Gly Asn Arg Val Arg
            180                 185                 190
Arg Ser Val Gly Ser Ser Leu Ser Cys Ile Asn Leu Asp Trp Asp Val
        195                 200                 205
Ile Arg Asp Lys Thr Lys Thr Lys Ile Glu Ser Leu Lys Glu His Gly
    210                 215                 220
Pro Ile Lys Asn Lys Met Ser Glu Ser Pro Asn Lys Thr Val Ser Glu
225                 230                 235                 240
```

```
Glu Lys Ala Lys Gln Tyr Leu Glu Glu Phe His Gln Thr Ala Leu Glu
                245                 250                 255
His Pro Glu Leu Ser Glu Leu Lys Thr Val Thr Gly Thr Asn Pro Val
            260                 265                 270
Phe Ala Gly Ala Asn Tyr Ala Ala Trp Ala Val Asn Val Ala Gln Val
        275                 280                 285
Ile Asp Ser Glu Thr Ala Asp Asn Leu Glu Lys Thr Thr Ala Ala Leu
    290                 295                 300
Ser Ile Leu Pro Gly Ile Gly Ser Val Met Gly Ile Ala Asp Gly Ala
305                 310                 315                 320
Val His His Asn Thr Glu Glu Ile Val Ala Gln Ser Ile Ala Leu Ser
                325                 330                 335
Ser Leu Met Val Ala Gln Ala Ile Pro Leu Val Gly Glu Leu Val Asp
            340                 345                 350
Ile Gly Phe Ala Ala Tyr Asn Phe Val Glu Ser Ile Ile Asn Leu Phe
        355                 360                 365
Gln Val Val His Asn Ser Tyr Asn Arg Pro Ala Tyr Ser Pro Gly His
    370                 375                 380
Lys Thr Gln Pro Phe Leu His Asp Gly Tyr Ala Val Ser Trp Asn Thr
385                 390                 395                 400
Val Glu Asp Ser Ile Ile Arg Thr Gly Phe Gln Gly Glu Ser Gly His
                405                 410                 415
Asp Ile Lys Ile Thr Ala Glu Asn Thr Pro Leu Pro Ile Ala Gly Val
            420                 425                 430
Leu Leu Pro Thr Ile Pro Gly Lys Leu Asp Val Asn Lys Ser Lys Thr
        435                 440                 445
His Ile Ser Val Asn Gly Arg Lys Ile Arg Met Arg Cys Arg Ala Ile
    450                 455                 460
Asp Gly Asp Val Thr Phe Cys Arg Pro Lys Ser Pro Val Tyr Val Gly
465                 470                 475                 480
Asn Gly Val His Ala Asn Leu His Val Ala Phe His Arg Ser Ser Ser
                485                 490                 495
Glu Lys Ile His Ser Asn Glu Ile Ser Ser Asp Ser Ile Gly Val Leu
            500                 505                 510
Gly Tyr Gln Lys Thr Val Asp His Thr Lys Val Asn Ser Lys Leu Ser
        515                 520                 525
Leu Phe Phe Glu Ile Lys Ser
    530                 535


<210> 789
<211> 164
<212> PRT
<213> Artificial Sequence

<220>
<223> Coat protein of Tobacco mosaic virus (TMV)

<400> 789
Met Met Ala Tyr Ser Ile Pro Thr Pro Ser Gln Leu Val Tyr Phe Thr
 1               5                  10                  15
Glu Asn Tyr Ala Asp Tyr Ile Pro Phe Val Asn Arg Leu Ile Asn Ala
            20                  25                  30
Arg Ser Asn Ser Phe Gln Thr Gln Ser Gly Arg Asp Glu Leu Arg Glu
        35                  40                  45
Ile Leu Ile Lys Ser Gln Val Ser Val Val Ser Pro Ile Ser Arg Phe
    50                  55                  60
Pro Ala Glu Pro Ala Tyr Tyr Ile Tyr Leu Arg Asp Pro Ser Ile Ser
65                  70                  75                  80
```

Thr Val Tyr Thr Ala Leu Leu Gln Ser Thr Asp Thr Arg Asn Arg Val
85 90 95
Ile Glu Val Glu Asn Ser Thr Asn Val Thr Thr Ala Glu Gln Leu Asn
100 105 110
Ala Val Arg Arg Thr Asp Asp Ala Ser Thr Ala Ile His Asn Asn Leu
115 120 125
Glu Gln Leu Leu Ser Leu Leu Thr Asn Gly Thr Gly Val Phe Asn Arg
130 135 140
Thr Ser Phe Glu Ser Ala Ser Gly Leu Thr Trp Leu Val Thr Thr Thr
145 150 155 160
Pro Arg Thr Ala

<210> 790
<211> 221
<212> PRT
<213> Artificial Sequence

<220>
<223> Coat protein of alfalfa mosaic virus (AMV)

<400> 790
Met Ser Ser Ser Gln Lys Lys Ala Gly Gly Lys Ala Gly Lys Pro Thr
1 5 10 15
Lys Arg Ser Gln Asn Tyr Ala Ala Leu Arg Lys Ala Gln Leu Pro Lys
20 25 30
Pro Pro Ala Leu Lys Val Pro Val Ala Lys Pro Thr Asn Thr Ile Leu
35 40 45
Pro Gln Thr Gly Cys Val Trp Gln Ser Leu Gly Thr Pro Leu Ser Leu
50 55 60
Ser Ser Ser Asn Gly Leu Gly Ala Arg Phe Leu Tyr Ser Phe Leu Lys
65 70 75 80
Asp Phe Ala Ala Pro Arg Ile Leu Glu Glu Asp Leu Ile Phe Arg Met
85 90 95
Val Phe Ser Ile Thr Pro Ser His Ala Gly Ser Phe Cys Leu Thr Asp
100 105 110
Asp Val Thr Thr Glu Asp Gly Arg Ala Val Ala His Gly Asn Pro Met
115 120 125
Gln Glu Phe Pro His Gly Ala Phe His Ala Asn Glu Lys Phe Gly Phe
130 135 140
Glu Leu Val Phe Thr Ala Pro Thr His Ala Gly Met Gln Asn Gln Asn
145 150 155 160
Phe Lys His Ser Tyr Ala Val Ala Leu Cys Leu Asp Phe Asp Ala Leu
165 170 175
Pro Glu Gly Ser Arg Asn Pro Ser Tyr Arg Phe Asn Glu Val Trp Val
180 185 190
Glu Arg Lys Ala Phe Pro Arg Ala Gly Pro Leu Arg Ser Leu Ile Thr
195 200 205
Val Gly Leu Phe Asp Asp Ala Asp Asp Leu Asp Arg Gln
210 215 220

<210> 791
<211> 237
<212> PRT
<213> Artificial Sequence

<220>

<223> Coat protein of Potato virus X

<400> 791
Met Thr Thr Pro Ala Asn Thr Thr Gln Ala Thr Gly Ser Thr Thr Ser
1 5 10 15
Thr Thr Thr Lys Thr Ala Gly Ala Thr Pro Ala Thr Thr Ser Gly Leu
20 25 30
Phe Thr Ile Pro Asp Gly Glu Phe Phe Ser Thr Ala Arg Ala Ile Val
35 40 45
Ala Ser Asn Ala Val Ala Thr Asn Glu Asp Leu Ser Lys Ile Glu Ala
50 55 60
Ile Trp Lys Asp Met Lys Val Pro Thr Asp Thr Met Ala Gln Ala Ala
65 70 75 80
Trp Asp Leu Val Arg His Cys Ala Asp Val Gly Ser Ser Ala Gln Thr
85 90 95
Glu Met Ile Asp Thr Gly Pro Tyr Ser Asn Gly Ile Ser Arg Ala Arg
100 105 110
Leu Ala Ala Ala Ile Lys Glu Val Cys Thr Leu Arg Gln Phe Cys Met
115 120 125
Lys Tyr Ala Pro Val Val Trp Asn Trp Met Leu Thr Asn Asn Ser Pro
130 135 140
Pro Ala Asn Trp Gln Ala Gln Gly Phe Lys Pro Glu His Lys Phe Ala
145 150 155 160
Ala Phe Asp Phe Phe Asn Gly Val Thr Asn Pro Ala Ala Ile Met Pro
165 170 175
Lys Glu Gly Leu Ile Arg Pro Pro Ser Glu Ala Glu Met Asn Ala Ala
180 185 190
Gln Thr Ala Ala Phe Val Lys Ile Thr Lys Ala Arg Ala Gln Ser Asn
195 200 205
Asp Phe Ala Ser Leu Asp Ala Ala Val Thr Arg Gly Arg Ile Thr Gly
210 215 220
Thr Thr Thr Ala Glu Ala Val Val Thr Leu Pro Pro Pro
225 230 235

<210> 792
<211> 395
<212> PRT
<213> Artificial Sequence

<220>
<223> class I outer membrane protein of Neisseria meningitidis

<400> 792
Met Arg Lys Lys Leu Thr Ala Leu Val Leu Ser Ala Leu Pro Leu Ala
1 5 10 15
Ala Val Ala Asp Val Ser Leu Tyr Gly Glu Ile Lys Ala Gly Val Glu
20 25 30
Gly Arg Asn Tyr Gln Leu Gln Leu Thr Glu Ala Gln Ala Ala Asn Gly
35 40 45
Gly Ala Ser Gly Gln Val Lys Val Thr Lys Val Thr Lys Ala Lys Ser
50 55 60
Arg Ile Arg Thr Lys Ile Ser Asp Phe Gly Ser Phe Ile Gly Phe Lys
65 70 75 80
Gly Ser Glu Asp Leu Gly Glu Gly Leu Lys Ala Val Trp Gln Leu Glu
85 90 95
Gln Asp Val Ser Val Ala Gly Gly Gly Ala Thr Gln Trp Gly Asn Arg
100 105 110

```
Glu Ser Phe Ile Gly Leu Ala Gly Glu Phe Gly Thr Leu Arg Ala Gly
        115                 120                 125
Arg Val Ala Asn Gln Phe Asp Asp Ala Ser Gln Ala Ile Asp Pro Trp
    130                 135                 140
Asp Ser Asn Asn Asp Val Ala Ser Gln Leu Gly Ile Phe Lys Arg His
145                 150                 155                 160
Asp Asp Met Pro Val Ser Val Arg Tyr Asp Ser Pro Glu Phe Ser Gly
                165                 170                 175
Phe Ser Gly Ser Val Gln Phe Val Pro Ala Gln Asn Ser Lys Ser Ala
            180                 185                 190
Tyr Lys Pro Ala Tyr Trp Thr Thr Val Asn Thr Gly Ser Ala Thr Thr
        195                 200                 205
Thr Thr Phe Val Pro Ala Val Val Gly Lys Pro Gly Ser Asp Val Tyr
    210                 215                 220
Tyr Ala Gly Leu Asn Tyr Lys Asn Gly Gly Phe Ala Gly Asn Tyr Ala
225                 230                 235                 240
Phe Lys Tyr Ala Arg His Ala Asn Val Gly Arg Asp Ala Phe Glu Leu
                245                 250                 255
Phe Leu Leu Gly Ser Gly Ser Asp Gln Ala Lys Gly Thr Asp Pro Leu
            260                 265                 270
Lys Asn His Gln Val His Arg Leu Thr Gly Gly Tyr Glu Glu Gly Gly
        275                 280                 285
Leu Asn Leu Ala Leu Ala Ala Gln Leu Asp Leu Ser Glu Asn Gly Asp
    290                 295                 300
Lys Thr Lys Asn Ser Thr Thr Glu Ile Ala Ala Thr Ala Ser Tyr Arg
305                 310                 315                 320
Phe Gly Asn Ala Val Pro Arg Ile Ser Tyr Ala His Gly Phe Asp Phe
                325                 330                 335
Ile Glu Arg Gly Lys Lys Gly Glu Asn Thr Ser Tyr Asp Gln Ile Ile
            340                 345                 350
Ala Gly Val Asp Tyr Asp Phe Ser Lys Arg Thr Ser Ala Ile Val Ser
        355                 360                 365
Gly Ala Trp Leu Lys Arg Asn Thr Gly Ile Gly Asn Tyr Thr Gln Ile
    370                 375                 380
Asn Ala Ala Ser Val Gly Leu Arg His Lys Phe
385                 390                 395


<210> 793
<211> 347
<212> PRT
<213> Artificial Sequence

<220>
<223> Major fimbrial subunit protein type I (Fimbrillin)

<400> 793
Met Val Leu Lys Thr Ser Asn Ser Asn Arg Ala Phe Gly Val Gly Asp
 1               5                  10                  15
Asp Glu Ser Lys Val Ala Lys Leu Thr Val Met Val Tyr Asn Gly Glu
            20                  25                  30
Gln Gln Glu Ala Ile Lys Ser Ala Glu Asn Ala Thr Lys Val Glu Asp
        35                  40                  45
Ile Lys Cys Ser Ala Gly Gln Arg Thr Leu Val Val Met Ala Asn Thr
    50                  55                  60
Gly Ala Met Glu Leu Val Gly Lys Thr Leu Ala Glu Val Lys Ala Leu
65                  70                  75                  80
Thr Thr Glu Leu Thr Ala Glu Asn Gln Glu Ala Ala Gly Leu Ile Met
                85                  90                  95
```

```
Thr Ala Glu Pro Lys Thr Ile Val Leu Lys Ala Gly Lys Asn Tyr Ile
            100                 105                 110
Gly Tyr Ser Gly Thr Gly Glu Gly Asn His Ile Glu Asn Asp Pro Leu
        115                 120                 125
Lys Ile Lys Arg Val His Ala Arg Met Ala Phe Thr Glu Ile Lys Val
    130                 135                 140
Gln Met Ser Ala Ala Tyr Asp Asn Ile Tyr Thr Phe Val Pro Glu Lys
145                 150                 155                 160
Ile Tyr Gly Leu Ile Ala Lys Lys Gln Ser Asn Leu Phe Gly Ala Thr
                165                 170                 175
Leu Val Asn Ala Asp Ala Asn Tyr Leu Thr Gly Ser Leu Thr Thr Phe
            180                 185                 190
Asn Gly Ala Tyr Thr Pro Ala Asn Tyr Ala Asn Val Pro Trp Leu Ser
        195                 200                 205
Arg Asn Tyr Val Ala Pro Ala Ala Asp Ala Pro Gln Gly Phe Tyr Val
    210                 215                 220
Leu Glu Asn Asp Tyr Ser Ala Asn Gly Gly Thr Ile His Pro Thr Ile
225                 230                 235                 240
Leu Cys Val Tyr Gly Lys Leu Gln Lys Asn Gly Ala Asp Leu Ala Gly
                245                 250                 255
Ala Asp Leu Ala Ala Ala Gln Ala Ala Asn Trp Val Asp Ala Glu Gly
            260                 265                 270
Lys Thr Tyr Tyr Pro Val Leu Val Asn Phe Asn Ser Asn Asn Tyr Thr
        275                 280                 285
Tyr Asp Ser Asn Tyr Thr Pro Lys Asn Lys Ile Glu Arg Asn His Lys
    290                 295                 300
Tyr Asp Ile Lys Leu Thr Ile Thr Gly Pro Gly Thr Asn Asn Pro Glu
305                 310                 315                 320
Asn Pro Ile Thr Glu Ser Ala His Leu Asn Val Gln Cys Thr Val Ala
                325                 330                 335
Glu Trp Val Leu Val Gly Gln Asn Ala Thr Trp
            340                 345
```

<210> 794
<211> 428
<212> PRT
<213> Artificial Sequence

<220>
<223> Mycoplasma fermentans macrophage activating lipopeptide (MALP-2)

<400> 794

```
Met Lys Lys Ser Lys Lys Ile Leu Leu Gly Leu Ser Pro Ile Ala Ala
 1                5                  10                  15
Val Leu Pro Ala Val Ala Val Ser Cys Gly Asn Asn Asp Glu Ser Asn
            20                  25                  30
Ile Ser Phe Lys Glu Lys Asp Ile Ser Lys Tyr Thr Thr Thr Asn Ala
        35                  40                  45
Asn Gly Lys Gln Val Val Lys Asn Ala Glu Leu Leu Lys Leu Lys Pro
    50                  55                  60
Val Leu Ile Thr Asp Glu Gly Lys Ile Asp Asp Lys Ser Phe Asn Gln
65                  70                  75                  80
Ser Ala Phe Glu Ala Leu Lys Ala Ile Asn Lys Gln Thr Gly Ile Glu
                85                  90                  95
Ile Asn Ser Val Glu Pro Ser Ser Asn Phe Glu Ser Ala Tyr Asn Ser
            100                 105                 110
Ala Leu Ser Ala Gly His Lys Ile Trp Val Leu Asn Gly Phe Lys His
```

```
        115                 120                 125
Gln Gln Ser Ile Lys Gln Tyr Ile Asp Ala His Arg Glu Glu Leu Glu
    130                 135                 140
Arg Asn Gln Ile Lys Ile Ile Gly Ile Asp Phe Asp Ile Glu Thr Glu
145                 150                 155                 160
Tyr Lys Trp Phe Tyr Ser Leu Gln Phe Asn Ile Lys Glu Ser Ala Phe
                165                 170                 175
Thr Thr Gly Tyr Ala Ile Ala Ser Trp Leu Ser Glu Gln Asp Glu Ser
            180                 185                 190
Lys Arg Val Val Ala Ser Phe Gly Val Gly Ala Phe Pro Gly Val Thr
        195                 200                 205
Thr Phe Asn Glu Gly Phe Ala Lys Gly Ile Leu Tyr Tyr Asn Gln Lys
    210                 215                 220
His Lys Ser Ser Lys Ile Tyr His Thr Ser Pro Val Lys Leu Asp Ser
225                 230                 235                 240
Gly Phe Thr Ala Gly Glu Lys Met Asn Thr Val Ile Asn Asn Val Leu
                245                 250                 255
Ser Ser Thr Pro Ala Asp Val Lys Tyr Asn Pro His Val Ile Leu Ser
            260                 265                 270
Val Ala Gly Pro Ala Thr Phe Glu Thr Val Arg Leu Ala Asn Lys Gly
        275                 280                 285
Gln Tyr Val Ile Gly Val Asp Ser Asp Gln Gly Met Ile Gln Asp Lys
    290                 295                 300
Asp Arg Ile Leu Thr Ser Val Leu Lys His Ile Lys Gln Ala Val Tyr
305                 310                 315                 320
Glu Thr Leu Leu Asp Leu Ile Leu Glu Lys Glu Glu Gly Tyr Lys Pro
                325                 330                 335
Tyr Val Val Lys Asp Lys Lys Ala Asp Lys Lys Trp Ser His Phe Gly
            340                 345                 350
Thr Gln Lys Glu Lys Trp Ile Gly Val Ala Glu Asn His Phe Ser Asn
        355                 360                 365
Thr Glu Glu Gln Ala Lys Ile Asn Asn Lys Ile Lys Glu Ala Ile Lys
    370                 375                 380
Met Phe Lys Glu Leu Pro Glu Asp Phe Val Lys Tyr Ile Asn Ser Asp
385                 390                 395                 400
Lys Ala Leu Lys Asp Gly Asn Lys Ile Asp Asn Val Ser Glu Arg Leu
                405                 410                 415
Glu Ala Ile Ile Ser Ala Ile Asn Lys Ala Ala Lys
            420                 425
```

<210> 795
<211> 143
<212> PRT
<213> Artificial Sequence

<220>
<223> p19 protein of Mycobacterium tuberculosis

<400> 795

```
Ala Thr Thr Leu Pro Val Gln Arg His Pro Arg Ser Leu Phe Pro Glu
 1               5                  10                  15
Phe Ser Glu Leu Phe Ala Ala Phe Pro Ser Phe Ala Gly Leu Arg Pro
            20                  25                  30
Thr Phe Asp Thr Arg Leu Met Arg Leu Glu Asp Glu Met Lys Glu Gly
        35                  40                  45
Arg Tyr Glu Val Arg Ala Glu Leu Pro Gly Val Asp Pro Asp Lys Asp
    50                  55                  60
Val Asp Ile Met Val Arg Asp Gly Gln Leu Thr Ile Lys Ala Glu Arg
```

65 70 75 80
Thr Glu Gln Lys Asp Phe Asp Gly Arg Ser Glu Phe Ala Tyr Gly Ser
85 90 95
Phe Val Arg Thr Val Ser Leu Pro Val Gly Ala Asp Glu Asp Asp Ile
100 105 110
Lys Ala Thr Tyr Asp Lys Gly Ile Leu Thr Val Ser Val Ala Val Ser
115 120 125
Glu Gly Lys Pro Thr Glu Lys His Ile Gln Ile Arg Ser Thr Asn
130 135 140

<210> 796
<211> 19
<212> PRT
<213> Artificial Sequence

<220>
<223> muturational cleavage site

<400> 796
Asn Val Pro Glu Lys Gln Thr Arg Gly Ile Phe Gly Ala Ile Ala Gly
1 5 10 15
Phe Ile Glu

<210> 797
<211> 19
<212> PRT
<213> Artificial Sequence

<220>
<223> muturational cleavage site

<400> 797
Asn Ile Pro Ser Ile Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly
1 5 10 15
Phe Ile Glu

<210> 798
<211> 19
<212> PRT
<213> Artificial Sequence

<220>
<223> muturational cleavage site

<400> 798
Pro Ala Lys Leu Leu Lys Glu Arg Gly Phe Phe Gly Ala Ile Ala Gly
1 5 10 15
Phe Leu Glu

<210> 799
<211> 19
<212> PRT

<213> Artificial Sequence

<220>
<223> muturational cleavage site

<400> 799
Arg Glu Arg Arg Arg Lys Lys Arg Gly Leu Phe Gly Ala Ile Ala Gly
1 5 10 15
Phe Ile Glu

<210> 800
<211> 19
<212> PRT
<213> Artificial Sequence

<220>
<223> muturational cleavage site

<400> 800
Asn Val Pro Glu Lys Gln Thr Arg Gly Ile Phe Gly Ala Ile Ala Gly
1 5 10 15
Phe Ile Glu

<210> 801
<211> 19
<212> PRT
<213> Artificial Sequence

<220>
<223> muturational cleavage site

<400> 801
Asn Val Pro Gln Ile Glu Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly
1 5 10 15
Phe Ile Glu

<210> 802
<211> 19
<212> PRT
<213> Artificial Sequence

<220>
<223> muturational cleavage site

<400> 802
Asn Ile Pro Ser Ile Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly
1 5 10 15
Phe Ile Glu

<210> 803
<211> 12

<212> PRT
<213> Artificial Sequence

<220>
<223> muturational cleavage site

<400> 803
Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu
1 5 10

<210> 804
<211> 19
<212> PRT
<213> Artificial Sequence

<220>
<223> muturational cleavage site

<400> 804
Pro Ala Lys Leu Leu Lys Glu Arg Gly Phe Phe Gly Ala Ile Ala Gly
1 5 10 15
Phe Leu Glu

<210> 805
<211> 12
<212> PRT
<213> Artificial Sequence

<220>
<223> muturational cleavage site

<400> 805
Arg Gly Phe Phe Gly Ala Ile Ala Gly Phe Leu Glu
1 5 10

<210> 806
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> muturational cleavage site

<400> 806
Ser Leu Leu Thr Glu Val Glu Thr Pro Ile Arg Asn Glu Trp Gly Ser
1 5 10 15
Arg Ser Asn Asp Ser Ser Asp Pro
20

<210> 807
<211> 29
<212> PRT
<213> Artificial Sequence

<220>
<223> muturational cleavage site

<400> 807
Gly Ser Gly Ala Gly Ser Leu Leu Thr Glu Val Glu Thr Pro Thr Arg
1 5 10 15
Asn Glu Trp Glu Cys Arg Cys Ser Asp Ser Ser Asp Pro
20 25

<210> 808
<211> 29
<212> PRT
<213> Artificial Sequence

<220>
<223> muturational cleavage site

<400> 808
Gly Ser Gly Ala Gly Ser Leu Leu Thr Glu Val Glu Thr Leu Thr Arg
1 5 10 15
Asn Gly Trp Gly Cys Arg Cys Ser Asp Ser Ser Asp Pro
20 25

<210> 809
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> muturational cleavage site

<400> 809
Leu Thr Ser Gly His Leu Lys Cys Arg Val Lys Met Glu Lys Leu Gln
1 5 10 15
Leu Lys Gly Thr
20

<210> 810
<211> 27
<212> PRT
<213> Artificial Sequence

<220>
<223> muturational cleavage site

<400> 810
Glu Ala Glu Pro Pro Phe Gly Asp Ser Tyr Ile Ile Ile Gly Val Glu
1 5 10 15
Pro Gly Gln Leu Lys Leu Asn Trp Phe Lys Lys
20 25

<210> 811
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> muturational cleavage site

<400> 811
Ser Leu Leu Thr Glu Val Glu Thr Pro Ile Arg Asn Glu Trp Gly Ser
1 5 10 15
Arg Ser Asn Asp Ser Ser Asp Pro
20

<210> 812
<211> 494
<212> PRT
<213> S. typhimurium flagellin (P06179)

<400> 812
Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn Asn
1 5 10 15
Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu Ser
20 25 30
Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln Ala
35 40 45
Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala Ser
50 55 60
Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly Ala
65 70 75 80
Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala Val
85 90 95
Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile Gln
100 105 110
Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly Gln
115 120 125
Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu Thr
130 135 140
Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu Lys
145 150 155 160
Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Thr Leu Asn Val Gln Gln
165 170 175
Lys Tyr Lys Val Ser Asp Thr Ala Ala Thr Val Thr Gly Tyr Ala Asp
180 185 190
Thr Thr Ile Ala Leu Asp Asn Ser Thr Phe Lys Ala Ser Ala Thr Gly
195 200 205
Leu Gly Gly Thr Asp Gln Lys Ile Asp Gly Asp Leu Lys Phe Asp Asp
210 215 220
Thr Thr Gly Lys Tyr Tyr Ala Lys Val Thr Val Thr Gly Gly Thr Gly
225 230 235 240
Lys Asp Gly Tyr Tyr Glu Val Ser Val Asp Lys Thr Asn Gly Glu Val
245 250 255
Thr Leu Ala Gly Gly Ala Thr Ser Pro Leu Thr Gly Gly Leu Pro Ala
260 265 270
Thr Ala Thr Glu Asp Val Lys Asn Val Gln Val Ala Asn Ala Asp Leu
275 280 285
Thr Glu Ala Lys Ala Ala Leu Thr Ala Ala Gly Val Thr Gly Thr Ala
290 295 300
Ser Val Val Lys Met Ser Tyr Thr Asp Asn Asn Gly Lys Thr Ile Asp
305 310 315 320
Gly Gly Leu Ala Val Lys Val Gly Asp Asp Tyr Tyr Ser Ala Thr Gln
325 330 335

```
Asn Lys Asp Gly Ser Ile Ser Ile Asn Thr Thr Lys Tyr Thr Ala Asp
            340                 345                 350
Asp Gly Thr Ser Lys Thr Ala Leu Asn Lys Leu Gly Gly Ala Asp Gly
        355                 360                 365
Lys Thr Glu Val Val Ser Ile Gly Gly Lys Thr Tyr Ala Ala Ser Lys
    370                 375                 380
Ala Glu Gly His Asn Phe Lys Ala Gln Pro Asp Leu Ala Glu Ala Ala
385                 390                 395                 400
Ala Thr Thr Thr Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala
                405                 410                 415
Gln Val Asp Thr Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe
            420                 425                 430
Asn Ser Ala Ile Thr Asn Leu Gly Asn Thr Val Asn Asn Leu Thr Ser
        435                 440                 445
Ala Arg Ser Arg Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn
    450                 455                 460
Met Ser Arg Ala Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala
465                 470                 475                 480
Gln Ala Asn Gln Val Pro Gln Asn Val Leu Ser Leu Leu Arg
                485                 490


<210> 813
<211> 504
<212> PRT
<213> Salmonella muenchen flagellin (P06177)

<400> 813
Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn Asn
 1               5                  10                  15
Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu Ser
            20                  25                  30
Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln Ala
        35                  40                  45
Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala Ser
    50                  55                  60
Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly Ala
65                  70                  75                  80
Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala Val
                85                  90                  95
Gln Ser Ala Asn Gly Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile Gln
            100                 105                 110
Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly Gln
        115                 120                 125
Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu Thr
    130                 135                 140
Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu Lys
145                 150                 155                 160
Glu Ile Ser Ser Lys Thr Leu Gly Leu Asp Lys Leu Asn Val Gln Asp
                165                 170                 175
Ala Tyr Thr Pro Lys Glu Thr Ala Val Thr Val Asp Lys Thr Thr Tyr
            180                 185                 190
Lys Asn Gly Thr Asp Thr Ile Thr Ala Gln Ser Asn Thr Asp Ile Gln
        195                 200                 205
Thr Ala Ile Gly Gly Gly Ala Thr Gly Val Thr Gly Ala Asp Ile Lys
    210                 215                 220
Phe Lys Asp Gly Gln Tyr Tyr Leu Asp Val Lys Gly Gly Ala Ser Ala
225                 230                 235                 240
Gly Val Tyr Lys Ala Thr Tyr Asp Glu Thr Thr Lys Lys Val Asn Ile
```

```
                245                 250                 255
Asp Thr Thr Asp Lys Thr Pro Leu Ala Thr Ala Glu Ala Thr Ala Ile
            260                 265                 270
Arg Gly Thr Ala Thr Ile Thr His Asn Gln Ile Ala Glu Val Thr Lys
        275                 280                 285
Glu Gly Val Asp Thr Thr Thr Val Ala Ala Gln Leu Ala Ala Ala Gly
    290                 295                 300
Val Thr Gly Ala Asp Lys Asp Asn Thr Ser Leu Val Lys Leu Ser Phe
305                 310                 315                 320
Glu Asp Lys Asn Gly Lys Val Ile Asp Gly Gly Tyr Ala Val Lys Met
                325                 330                 335
Gly Asp Asp Phe Tyr Ala Ala Thr Tyr Asp Glu Lys Gln Val Gln Leu
            340                 345                 350
Leu Leu Asn Asn His Tyr Thr Asp Gly Ala Gly Val Leu Gln Thr Gly
        355                 360                 365
Ala Val Lys Phe Gly Gly Ala Asn Gly Lys Ser Glu Val Val Thr Ala
    370                 375                 380
Thr Val Gly Lys Thr Tyr Leu Ala Ser Asp Leu Asp Lys His Asn Phe
385                 390                 395                 400
Arg Thr Gly Gly Glu Leu Lys Glu Val Asn Thr Asp Lys Thr Glu Asn
                405                 410                 415
Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Thr Leu Arg
            420                 425                 430
Ser Asp Leu Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile Thr Asn
        435                 440                 445
Leu Gly Asn Thr Val Asn Asn Leu Ser Ser Ala Arg Ser Arg Ile Glu
    450                 455                 460
Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala Gln Ile
465                 470                 475                 480
Leu Gln Gln Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln Val Pro
                485                 490                 495
Gln Asn Val Leu Ser Leu Leu Arg
            500
```

<210> 814
<211> 497
<212> PRT
<213> E. coli flagellin (P04949)

<400> 814

```
Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Ile Thr Gln Asn Asn
 1               5                  10                  15
Ile Asn Lys Asn Gln Ser Ala Leu Ser Ser Ser Ile Glu Arg Leu Ser
            20                  25                  30
Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln Ala
        35                  40                  45
Ile Ala Asn Arg Phe Thr Ser Asn Ile Lys Gly Leu Thr Gln Ala Ala
    50                  55                  60
Arg Asn Ala Asn Asp Gly Ile Ser Val Ala Gln Thr Thr Glu Gly Ala
65                  70                  75                  80
Leu Ser Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Thr Val
                85                  90                  95
Gln Ala Thr Thr Gly Thr Asn Ser Glu Ser Asp Leu Ser Ser Ile Gln
            100                 105                 110
Asp Glu Ile Lys Ser Arg Leu Asp Glu Ile Asp Arg Val Ser Gly Gln
        115                 120                 125
Thr Gln Phe Asn Gly Val Asn Val Leu Ala Lys Asn Gly Ser Met Lys
    130                 135                 140
```

```
Ile Gln Val Gly Ala Asn Asp Asn Gln Thr Ile Thr Ile Asp Leu Lys
145                 150                 155                 160
Gln Ile Asp Ala Lys Thr Leu Gly Leu Asp Gly Phe Ser Val Lys Asn
                165                 170                 175
Asn Asp Thr Val Thr Thr Ser Ala Pro Val Thr Ala Phe Gly Ala Thr
            180                 185                 190
Thr Thr Asn Asn Ile Lys Leu Thr Gly Ile Thr Leu Ser Thr Glu Ala
        195                 200                 205
Ala Thr Asp Thr Gly Gly Thr Asn Pro Ala Ser Ile Glu Gly Val Tyr
    210                 215                 220
Thr Asp Asn Gly Asn Asp Tyr Tyr Ala Lys Ile Thr Gly Gly Asp Asn
225                 230                 235                 240
Asp Gly Lys Tyr Tyr Ala Val Thr Val Ala Asn Asp Gly Thr Val Thr
                245                 250                 255
Met Ala Thr Gly Ala Thr Ala Asn Ala Thr Val Thr Asp Ala Asn Thr
            260                 265                 270
Thr Lys Ala Thr Thr Ile Thr Ser Gly Gly Thr Pro Val Gln Ile Asp
        275                 280                 285
Asn Thr Ala Gly Ser Ala Thr Ala Asn Leu Gly Ala Val Ser Leu Val
    290                 295                 300
Lys Leu Gln Asp Ser Lys Gly Asn Asp Thr Asp Thr Tyr Ala Leu Lys
305                 310                 315                 320
Asp Thr Asn Gly Asn Leu Tyr Ala Ala Asp Val Asn Glu Thr Thr Gly
                325                 330                 335
Ala Val Ser Val Lys Thr Ile Thr Tyr Thr Asp Ser Ser Gly Ala Ala
            340                 345                 350
Ser Ser Pro Thr Ala Val Lys Leu Gly Gly Asp Asp Gly Lys Thr Glu
        355                 360                 365
Val Val Asp Ile Asp Gly Lys Thr Tyr Asp Ser Ala Asp Leu Asn Gly
    370                 375                 380
Gly Asn Leu Gln Thr Gly Leu Thr Ala Gly Gly Glu Ala Leu Thr Ala
385                 390                 395                 400
Val Ala Asn Gly Lys Thr Thr Asp Pro Leu Lys Ala Leu Asp Asp Ala
                405                 410                 415
Ile Ala Ser Val Asp Lys Phe Arg Ser Ser Leu Gly Ala Val Gln Asn
            420                 425                 430
Arg Leu Asp Ser Ala Val Thr Asn Leu Asn Asn Thr Thr Thr Asn Leu
        435                 440                 445
Ser Glu Ala Gln Ser Arg Ile Gln Asp Ala Asp Tyr Ala Thr Glu Val
    450                 455                 460
Ser Asn Met Ser Lys Ala Gln Ile Ile Gln Gln Ala Gly Asn Ser Val
465                 470                 475                 480
Leu Ala Lys Ala Asn Gln Val Pro Gln Gln Val Leu Ser Leu Leu Gln
                485                 490                 495
Gly
```

<210> 815
<211> 393
<212> PRT
<213> pseudomonas aerugienosa

```
<400> 815
Ala Leu Thr Val Asn Thr Asn Ile Ala Ser Leu Asn Thr Gln Arg Asn
 1               5                  10                  15
Leu Asn Asn Ser Ser Ala Ser Leu Asn Thr Ser Leu Gln Arg Leu Ser
            20                  25                  30
Thr Gly Ser Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Leu Gln
```

```
        35                  40                  45
Ile Ala Asn Arg Leu Thr Ser Gln Val Asn Gly Leu Asn Val Ala Thr
    50                  55                  60
Lys Asn Ala Asn Asp Gly Ile Ser Leu Ala Gln Thr Ala Glu Gly Ala
65                  70                  75                  80
Leu Gln Gln Ser Thr Asn Ile Leu Gln Arg Met Arg Asp Leu Ser Leu
                85                  90                  95
Gln Ser Ala Asn Gly Ser Asn Ser Asp Ser Glu Arg Thr Ala Leu Asn
            100                 105                 110
Gly Glu Val Lys Gln Leu Gln Lys Glu Leu Asp Arg Ile Ser Asn Thr
        115                 120                 125
Thr Thr Phe Gly Gly Arg Lys Leu Leu Asp Gly Ser Phe Gly Val Ala
    130                 135                 140
Ser Phe Gln Val Gly Ser Ala Ala Asn Glu Ile Ile Ser Val Gly Ile
145                 150                 155                 160
Asp Glu Met Ser Ala Glu Ser Leu Asn Gly Thr Tyr Phe Lys Ala Asp
                165                 170                 175
Gly Gly Gly Ala Val Thr Ala Ala Thr Ala Ser Gly Thr Val Asp Ile
            180                 185                 190
Ala Ile Gly Ile Thr Gly Gly Ser Ala Val Asn Val Lys Val Asp Met
        195                 200                 205
Lys Gly Asn Glu Thr Ala Glu Gln Ala Ala Ala Lys Ile Ala Ala Ala
    210                 215                 220
Val Asn Asp Ala Asn Val Gly Ile Gly Ala Phe Ser Asp Gly Asp Thr
225                 230                 235                 240
Ile Ser Tyr Val Ser Lys Ala Gly Lys Asp Gly Ser Gly Ala Ile Thr
                245                 250                 255
Ser Ala Val Ser Gly Val Val Ile Ala Asp Thr Gly Ser Thr Gly Val
            260                 265                 270
Gly Thr Ala Ala Gly Val Thr Pro Ser Ala Thr Ala Phe Ala Lys Thr
        275                 280                 285
Asn Asp Thr Val Ala Lys Ile Asp Ile Ser Thr Ala Lys Gly Ala Gln
    290                 295                 300
Ser Ala Val Leu Val Ile Asp Glu Ala Ile Lys Gln Ile Asp Ala Gln
305                 310                 315                 320
Arg Ala Asp Leu Gly Ala Val Gln Asn Arg Phe Asp Asn Thr Ile Asn
                325                 330                 335
Asn Leu Lys Asn Ile Gly Glu Asn Val Ser Ala Ala Arg Gly Arg Ile
            340                 345                 350
Glu Asp Thr Asp Phe Ala Ala Glu Thr Ala Asn Leu Thr Lys Asn Gln
        355                 360                 365
Val Leu Gln Gln Ala Gly Thr Ala Ile Leu Ala Gln Ala Asn Gln Leu
    370                 375                 380
Pro Gln Ser Val Leu Ser Leu Leu Arg
385                 390


<210> 816
<211> 506
<212> PRT
<213> Artificial Sequence

<220>
<223> STF2

<400> 816
Met Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn
 1               5                  10                  15
Asn Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu
```

```
            20                  25                  30
Ser Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln
        35                  40                  45
Ala Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala
    50                  55                  60
Ser Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly
65                  70                  75                  80
Ala Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala
                85                  90                  95
Val Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile
            100                 105                 110
Gln Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly
        115                 120                 125
Gln Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu
    130                 135                 140
Thr Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu
145                 150                 155                 160
Lys Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val Gln
                165                 170                 175
Lys Ala Tyr Asp Val Lys Asp Thr Ala Val Thr Thr Lys Ala Tyr Ala
            180                 185                 190
Asn Asn Gly Thr Thr Leu Asp Val Ser Gly Leu Asp Asp Ala Ala Ile
        195                 200                 205
Lys Ala Ala Thr Gly Gly Thr Asn Gly Thr Ala Ser Val Thr Gly Gly
    210                 215                 220
Ala Val Lys Phe Asp Ala Asp Asn Asn Lys Tyr Phe Val Thr Ile Gly
225                 230                 235                 240
Gly Phe Thr Gly Ala Asp Ala Ala Lys Asn Gly Asp Tyr Glu Val Asn
                245                 250                 255
Val Ala Thr Asp Gly Thr Val Thr Leu Ala Ala Gly Ala Thr Lys Thr
            260                 265                 270
Thr Met Pro Ala Gly Ala Thr Thr Lys Thr Glu Val Gln Glu Leu Lys
        275                 280                 285
Asp Thr Pro Ala Val Val Ser Ala Asp Ala Lys Asn Ala Leu Ile Ala
    290                 295                 300
Gly Gly Val Asp Ala Thr Asp Ala Asn Gly Ala Glu Leu Val Lys Met
305                 310                 315                 320
Ser Tyr Thr Asp Lys Asn Gly Lys Thr Ile Glu Gly Gly Tyr Ala Leu
                325                 330                 335
Lys Ala Gly Asp Lys Tyr Tyr Ala Ala Asp Tyr Asp Glu Ala Thr Gly
            340                 345                 350
Ala Ile Lys Ala Lys Thr Thr Ser Tyr Thr Ala Ala Asp Gly Thr Thr
        355                 360                 365
Lys Thr Ala Ala Asn Gln Leu Gly Gly Val Asp Gly Lys Thr Glu Val
    370                 375                 380
Val Thr Ile Asp Gly Lys Thr Tyr Asn Ala Ser Lys Ala Ala Gly His
385                 390                 395                 400
Asp Phe Lys Ala Gln Pro Glu Leu Ala Glu Ala Ala Ala Lys Thr Thr
                405                 410                 415
Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala
            420                 425                 430
Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile
        435                 440                 445
Thr Asn Leu Gly Asn Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg
    450                 455                 460
Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala
465                 470                 475                 480
Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln
                485                 490                 495
```

Val Pro Gln Asn Val Leu Ser Leu Leu Arg
500 505

<210> 817
<211> 345
<212> DNA
<213> Artificial Sequence

<220>
<223> EIII+

<400> 817
atggaaaaat tgcagttgaa gggaacaacc tatggcgtct gttcaaaggc tttcaagttt 60
cttgggactc ccgcagacac aggtcacggc actgtggtgt tggaattgca gtacactggc 120
acggatggac cttgcaaagt tcctatctcg tcagtggctt cattgaacga cctaacgcca 180
gtgggcagat tggtcactgt caacccttt gtttcagtgg ccacggccaa cgctaaggtc 240
ctgattgaat tggaaccacc ctttggagac tcatacatag tggtgggcag aggagaacaa 300
cagatcaatc accattggca caagtctgga agcagcattg gcaaa 345

<210> 818
<211> 116
<212> PRT
<213> Artificial Sequence

<220>
<223> JEIII+

<400> 818
Met Asp Lys Leu Ala Leu Lys Gly Thr Thr Tyr Gly Met Cys Thr Glu
1 5 10 15
Lys Phe Ser Phe Ala Lys Asn Pro Val Asp Thr Gly His Gly Thr Val
20 25 30
Val Ile Glu Leu Ser Tyr Ser Gly Ser Asp Gly Pro Cys Lys Ile Pro
35 40 45
Ile Val Ser Val Ala Ser Leu Asn Asp Met Thr Pro Val Gly Arg Leu
50 55 60
Val Thr Val Asn Pro Phe Val Ala Thr Ser Ser Ala Asn Ser Lys Val
65 70 75 80
Leu Val Glu Met Glu Pro Pro Phe Gly Asp Ser Tyr Ile Val Val Gly
85 90 95
Arg Gly Asp Lys Gln Ile Asn His His Trp His Lys Ala Gly Ser Thr
100 105 110
Leu Gly Lys Ala
115

<210> 819
<211> 348
<212> DNA
<213> Artificial Sequence

<220>
<223> JEIII+

<400> 819
atggacaaac tggctctgaa aggcacaacc tatggcatgt gtacagaaaa attctcgttc 60
gcgaaaaatc cggtggacac tggtcacgga acagttgtca ttgaactctc ctactctggg 120
agtgatggcc cctgcaaaat tccgattgtt tccgttgcga gcctcaatga catgaccccc 180

gttgggcggc tggtgacagt gaacccttc gtcgcgactt ccagtgccaa ctcaaaggtg 240
ctggtcgaga tggaaccccc cttcggagac tcctacatcg tagttggaag gggagacaag 300
cagatcaacc accattggca caaagctgga agcacgctgg gcaaggcc 348

<210> 820
<211> 287
<212> PRT
<213> Listeria monocytogenes

<400> 820
Met Lys Val Asn Thr Asn Ile Ile Ser Leu Lys Thr Gln Glu Tyr Leu
1 5 10 15
Arg Lys Asn Asn Glu Gly Met Thr Gln Ala Gln Glu Arg Leu Ala Ser
20 25 30
Gly Lys Arg Ile Asn Ser Ser Leu Asp Asp Ala Ala Gly Leu Ala Val
35 40 45
Val Thr Arg Met Asn Val Lys Ser Thr Gly Leu Asp Ala Ala Ser Lys
50 55 60
Asn Ser Ser Met Gly Ile Asp Leu Leu Gln Thr Ala Asp Ser Ala Leu
65 70 75 80
Ser Ser Met Ser Ser Ile Leu Gln Arg Met Arg Gln Leu Ala Val Gln
85 90 95
Ser Ser Asn Gly Ser Phe Ser Asp Glu Asp Arg Lys Gln Tyr Thr Ala
100 105 110
Glu Phe Gly Ser Leu Ile Lys Glu Leu Asp His Val Ala Asp Thr Thr
115 120 125
Asn Tyr Asn Asn Ile Lys Leu Leu Asp Gln Thr Ala Thr Asn Ala Ala
130 135 140
Thr Gln Val Ser Ile Gln Ala Ser Asp Lys Ala Asn Asp Leu Ile Asn
145 150 155 160
Ile Asp Leu Phe Asn Ala Lys Gly Leu Ser Ala Gly Thr Ile Thr Leu
165 170 175
Gly Ser Gly Ser Thr Val Ala Gly Tyr Ser Ala Leu Ser Val Ala Asp
180 185 190
Ala Asp Ser Ser Gln Glu Ala Thr Glu Ala Ile Asp Glu Leu Ile Asn
195 200 205
Asn Ile Ser Asn Gly Arg Ala Leu Leu Gly Ala Gly Met Ser Arg Leu
210 215 220
Ser Tyr Asn Val Ser Asn Val Asn Asn Gln Ser Ile Ala Thr Lys Ala
225 230 235 240
Ser Ala Ser Ser Ile Glu Asp Ala Asp Met Ala Ala Glu Met Ser Glu
245 250 255
Met Thr Lys Tyr Lys Ile Leu Thr Gln Thr Ser Ile Ser Met Leu Ser
260 265 270
Gln Ala Asn Gln Thr Pro Gln Met Leu Thr Gln Leu Ile Asn Ser
275 280 285

<210> 821
<211> 12
<212> PRT
<213> Artificial Sequence

<220>
<223> muturational cleavage site

<220>
<221> VARIANT
<222> 4

<223> Xaa = Any Amino Acid

<400> 821
Arg Gly Leu Xaa Gly Ala Ile Ala Gly Phe Ile Glu
1 5 10

<210> 822
<211> 12
<212> PRT
<213> Artificial Sequence

<220>
<223> muturational cleavage site

<220>
<221> VARIANT
<222> 4
<223> Xaa = Any Amino Acid

<400> 822
Arg Gly Leu Xaa Gly Ala Ile Ala Gly Phe Ile Glu
1 5 10

<210> 823
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> flagellin peptide

<400> 823
Thr Ile Ala Leu
1

<210> 824
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> flagellin peptide

<400> 824
Thr Thr Leu Asp
1

<210> 825
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> flagellin peptide

<400> 825
Gly Thr Asp Gln Lys Ile Asp
1 5

<210> 826
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> flagellin peptide

<400> 826
Asn Gly Glu Val Thr Leu
1 5

<210> 827
<211> 12
<212> PRT
<213> Artificial Sequence

<220>
<223> muturational cleavage site

<400> 827
Ser Gly Leu Asp Asp Ala Ala Ile Lys Ala Ala Thr
1 5 10

<210> 828
<211> 12
<212> PRT
<213> Artificial Sequence

<220>
<223> muturational cleavage site

<400> 828
Arg Gly Ile Phe Gly Ala Ile Ala Gly Phe Ile Glu
1 5 10

<210> 829
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> flagellin peptides

<400> 829
Gly Ala Asp Ala Ala
1 5

<210> 830
<211> 4

<212> PRT
<213> Artificial Sequence

<220>
<223> flagellin peptides

<400> 830
Ala Gly Gly Ala
1

<210> 831
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> flagellin peptides

<400> 831
Pro Ala Thr Ala
1

<210> 832
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> flagellin peptides

<400> 832
Ala Ala Gly Ala
1

<210> 833
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> flagellin peptides

<400> 833
Ala Thr Thr Lys
1

<210> 834
<211> 12
<212> PRT
<213> Artificial Sequence

<220>
<223> flagellin peptides

<400> 834

Ala Gly Ala Thr Lys Thr Thr Met Pro Ala Gly Ala
1 5 10

<210> 835
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> flagellin peptides

<400> 835
Val Thr Gly Thr Gly
1 5

<210> 836
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> flagellin peptides

<400> 836
Thr Glu Ala Lys Ala Ala Leu Thr Ala Ala
1 5 10

<210> 837
<211> 11
<212> PRT
<213> Artificial Sequence

<220>
<223> flagellin peptides

<400> 837
Ala Ser Val Val Lys Met Ser Tyr Thr Asp Asn
1 5 10

<210> 838
<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> flagellin peptides

<400> 838
Val Asp Ala Thr Asp Ala Asn Gly Ala
1 5

<210> 839
<211> 19
<212> PRT

<213> Artificial Sequence

<220>
<223> muturational cleavage site

<400> 839
Arg Glu Arg Arg Arg Lys Lys Arg Gly Leu Phe Gly Ala Ile Ala Gly
1 5 10 15
Phe Ile Glu

<210> 840
<211> 494
<212> PRT
<213> Artificial Sequence

<220>
<223> flagellin 1UCU

<400> 840
Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn Asn
1 5 10 15
Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu Ser
20 25 30
Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln Ala
35 40 45
Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala Ser
50 55 60
Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly Ala
65 70 75 80
Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala Val
85 90 95
Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile Gln
100 105 110
Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly Gln
115 120 125
Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu Thr
130 135 140
Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu Lys
145 150 155 160
Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Thr Leu Asn Val Gln Gln
165 170 175
Lys Tyr Lys Val Ser Asp Thr Ala Ala Thr Val Thr Gly Tyr Ala Asp
180 185 190
Thr Thr Ile Ala Leu Asp Asn Ser Thr Phe Lys Ala Ser Ala Thr Gly
195 200 205
Leu Gly Gly Thr Asp Gln Lys Ile Asp Gly Asp Leu Lys Phe Asp Asp
210 215 220
Thr Thr Gly Lys Tyr Tyr Ala Lys Val Thr Val Thr Gly Gly Thr Gly
225 230 235 240
Lys Asp Gly Tyr Tyr Glu Val Ser Val Asp Lys Thr Asn Gly Glu Val
245 250 255
Thr Leu Ala Gly Gly Ala Thr Ser Pro Leu Thr Gly Gly Leu Pro Ala
260 265 270
Thr Ala Thr Glu Asp Val Lys Asn Val Gln Val Ala Asn Ala Asp Leu
275 280 285
Thr Glu Ala Lys Ala Ala Leu Thr Ala Ala Gly Val Thr Gly Thr Ala
290 295 300

```
Ser Val Val Lys Met Ser Tyr Thr Asp Asn Asn Gly Lys Thr Ile Asp
305                 310                 315                 320
Gly Gly Leu Ala Val Lys Val Gly Asp Asp Tyr Tyr Ser Ala Thr Gln
                325                 330                 335
Asn Lys Asp Gly Ser Ile Ser Ile Asn Thr Thr Lys Tyr Thr Ala Asp
            340                 345                 350
Asp Gly Thr Ser Lys Thr Ala Leu Asn Lys Leu Gly Gly Ala Asp Gly
        355                 360                 365
Lys Thr Glu Val Val Ser Ile Gly Gly Lys Thr Tyr Ala Ala Ser Lys
    370                 375                 380
Ala Glu Gly His Asn Phe Lys Ala Gln Pro Asp Leu Ala Glu Ala Ala
385                 390                 395                 400
Ala Thr Thr Thr Glu Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala
                405                 410                 415
Gln Val Asp Thr Leu Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe
            420                 425                 430
Asn Ser Ala Ile Thr Asn Leu Gly Asn Thr Val Asn Asn Leu Thr Ser
        435                 440                 445
Val Arg Ser Arg Ile Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn
    450                 455                 460
Met Ser Arg Ala Gln Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala
465                 470                 475                 480
Gln Ala Asn Gln Val Pro Gln Asn Val Leu Ser Leu Leu Arg
                485                 490
```

<210> 841
<211> 505
<212> PRT
<213> Artificial Sequence

<220>
<223> flagellin Sf1a

```
<400> 841
Ala Gln Val Ile Asn Thr Asn Ser Leu Ser Leu Leu Thr Gln Asn Asn
 1               5                  10                  15
Leu Asn Lys Ser Gln Ser Ala Leu Gly Thr Ala Ile Glu Arg Leu Ser
            20                  25                  30
Ser Gly Leu Arg Ile Asn Ser Ala Lys Asp Asp Ala Ala Gly Gln Ala
        35                  40                  45
Ile Ala Asn Arg Phe Thr Ala Asn Ile Lys Gly Leu Thr Gln Ala Ser
    50                  55                  60
Arg Asn Ala Asn Asp Gly Ile Ser Ile Ala Gln Thr Thr Glu Gly Ala
65                  70                  75                  80
Leu Asn Glu Ile Asn Asn Asn Leu Gln Arg Val Arg Glu Leu Ala Val
                85                  90                  95
Gln Ser Ala Asn Ser Thr Asn Ser Gln Ser Asp Leu Asp Ser Ile Gln
            100                 105                 110
Ala Glu Ile Thr Gln Arg Leu Asn Glu Ile Asp Arg Val Ser Gly Gln
        115                 120                 125
Thr Gln Phe Asn Gly Val Lys Val Leu Ala Gln Asp Asn Thr Leu Thr
    130                 135                 140
Ile Gln Val Gly Ala Asn Asp Gly Glu Thr Ile Asp Ile Asp Leu Lys
145                 150                 155                 160
Gln Ile Asn Ser Gln Thr Leu Gly Leu Asp Ser Leu Asn Val Gln Lys
                165                 170                 175
Ala Tyr Asp Val Lys Asp Thr Ala Val Thr Thr Lys Ala Tyr Ala Asn
            180                 185                 190
```

```
Asn Gly Thr Thr Leu Asp Val Ser Gly Leu Asp Asp Ala Ala Ile Lys
        195                 200                 205
Ala Ala Thr Gly Gly Thr Asn Gly Thr Ala Ser Val Thr Gly Gly Ala
    210                 215                 220
Val Lys Phe Asp Ala Asp Asn Asn Lys Tyr Phe Val Thr Ile Gly Gly
225                 230                 235                 240
Phe Thr Gly Ala Asp Ala Ala Lys Asn Gly Asp Tyr Glu Val Asn Val
                245                 250                 255
Ala Thr Asp Gly Thr Val Thr Leu Ala Ala Gly Ala Thr Lys Thr Thr
            260                 265                 270
Met Pro Ala Gly Ala Thr Thr Lys Thr Glu Val Gln Glu Leu Lys Asp
        275                 280                 285
Thr Pro Ala Val Val Ser Ala Asp Ala Lys Asn Ala Leu Ile Ala Gly
    290                 295                 300
Gly Val Asp Ala Thr Asp Ala Asn Gly Ala Glu Leu Val Lys Met Ser
305                 310                 315                 320
Tyr Thr Asp Lys Asn Gly Lys Thr Ile Glu Gly Gly Tyr Ala Leu Lys
                325                 330                 335
Ala Gly Asp Lys Tyr Tyr Ala Ala Asp Tyr Asp Glu Ala Thr Gly Ala
            340                 345                 350
Ile Lys Ala Lys Thr Thr Ser Tyr Thr Ala Ala Asp Gly Thr Thr Lys
        355                 360                 365
Thr Ala Ala Asn Gln Leu Gly Gly Val Asp Gly Lys Thr Glu Val Val
    370                 375                 380
Thr Ile Asp Gly Lys Thr Tyr Asn Ala Ser Lys Ala Ala Gly His Asp
385                 390                 395                 400
Phe Lys Ala Gln Pro Glu Leu Ala Glu Ala Ala Ala Lys Thr Thr Glu
                405                 410                 415
Asn Pro Leu Gln Lys Ile Asp Ala Ala Leu Ala Gln Val Asp Ala Leu
            420                 425                 430
Arg Ser Asp Leu Gly Ala Val Gln Asn Arg Phe Asn Ser Ala Ile Thr
        435                 440                 445
Asn Leu Gly Asn Thr Val Asn Asn Leu Ser Glu Ala Arg Ser Arg Ile
    450                 455                 460
Glu Asp Ser Asp Tyr Ala Thr Glu Val Ser Asn Met Ser Arg Ala Gln
465                 470                 475                 480
Ile Leu Gln Gln Ala Gly Thr Ser Val Leu Ala Gln Ala Asn Gln Val
                485                 490                 495
Pro Gln Asn Val Leu Ser Leu Leu Arg
            500                 505
```