Innovation, Sciences et Développement économique Canada
Office de la Propriété Intellectuelle du Canada

Innovation, Science and Economic Development Canada
Canadian Intellectual Property Office

CA 3092925 A1 2019/09/06

(21) 3 092 925

(12) DEMANDE DE BREVET CANADIEN
CANADIAN PATENT APPLICATION

(13) A1

(86) Date de dépôt PCT/PCT Filing Date: 2019/03/01
(87) Date publication PCT/PCT Publication Date: 2019/09/06
(85) Entrée phase nationale/National Entry: 2020/09/02
(86) N° demande PCT/PCT Application No.: US 2019/020436
(87) N° publication PCT/PCT Publication No.: 2019/169356
(30) Priorités/Priorities: 2018/03/02 (US PCT/US18/20788); 2018/10/01 (US62/739,701); 2018/10/19 (US62/748,292)

(51) Cl.Int./Int.Cl. *A61K 38/16* (2006.01), *A61K 39/12* (2006.01), *A61K 39/21* (2006.01), *A61K 39/39* (2006.01)

(71) Demandeur/Applicant:
DUKE UNIVERSITY, US

(72) Inventeur/Inventor:
SAUNDERS, KEVIN, US

(74) Agent: GOWLING WLG (CANADA) LLP

(54) Titre : COMPOSITIONS COMPRENANT DES ENVELOPPES DE VIH POUR INDUIRE DES ANTICORPS ANTI-VIH -1

(54) Title: COMPOSITIONS COMPRISING HIV ENVELOPES TO INDUCE HIV-1 ANTIBODIES

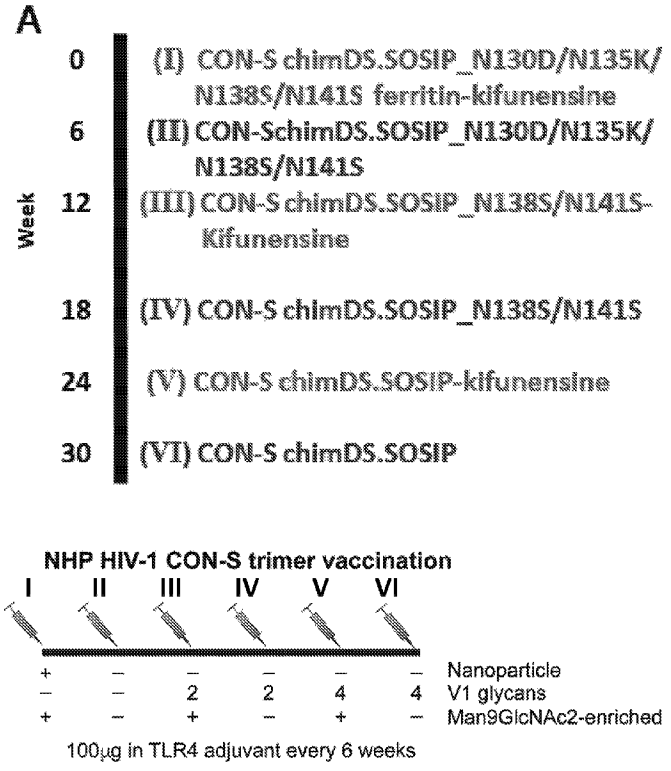


Figure 3A

(57) Abrégé/Abstract:
The invention is directed to modified HIV-1 envelopes, compositions comprising these modified envelopes and methods of using these modified HIV-1 envelopes to induce immune responses.

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
06 September 2019 (06.09.2019)

WIPO | PCT

(10) International Publication Number
WO 2019/169356 A1

(51) International Patent Classification:
A61K 38/16 (2006.01) *A61K 39/21* (2006.01)
A61K 39/12 (2006.01) *A61K 39/39* (2006.01)

(21) International Application Number:
PCT/US2019/020436

(22) International Filing Date:
01 March 2019 (01.03.2019)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

PCT/US18/20788	02 March 2018 (02.03.2018)	US
62/739,701	01 October 2018 (01.10.2018)	US
62/748,292	19 October 2018 (19.10.2018)	US

(71) Applicant: **DUKE UNIVERSITY** [US/US]; 2812 Erwin Road, Suite 306, Durham, NC 27705 (US).

(72) Inventor: **SAUNDERS, Kevin**; c/o Duke University, 2812 Erwin Road, Suite 306, Durham, NC 27705 (US).

(74) Agent: **KIM, William, W.** et al.; Wilmer Cutler Pickering Hale and Dorr LLP, 1875 Pennsylvania Avenue, NW, Washington, DC 20006 (US).

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

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

(54) Title: COMPOSITIONS COMPRISING HIV ENVELOPES TO INDUCE HIV-1 ANTIBODIES

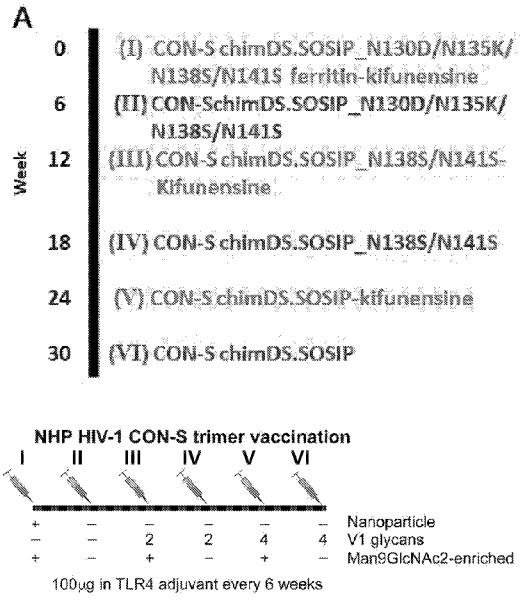


Figure 3A

(57) Abstract: The invention is directed to modified HIV-1 envelopes, compositions comprising these modified envelopes and methods of using these modified HIV-1 envelopes to induce immune responses.

WO 2019/169356 A1

[Continued on next page]

WO 2019/169356 A1

Published:

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

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

Compositions comprising HIV envelopes to induce HIV-1 antibodies

[0001] This application claims the benefit and priority to International PCT application PCT/US2018/020788 filed March 2, 2018, U.S. Application Serial Number 62/739,701 filed October 1, 2018 and U.S. Application Serial Number 62/748,292 filed October 19, 2018, the entire contents of each application are herein incorporated by reference.

[0002] This invention was made with government support under NIAID Research Grant (R01AI120801). The government has certain rights in the invention.

TECHNICAL FIELD

[0003] The present invention relates in general, to a composition suitable for use in inducing anti-HIV-1 antibodies, and, in particular, to immunogenic compositions comprising envelope proteins and nucleic acids to induce cross-reactive neutralizing antibodies and increase their breadth of coverage. The invention also relates to methods of inducing such broadly neutralizing anti-HIV-1 antibodies using such compositions.

BACKGROUND

[0004] The development of a safe and effective HIV-1 vaccine is one of the highest priorities of the scientific community working on the HIV-1 epidemic. While anti-retroviral treatment (ART) has dramatically prolonged the lives of HIV-1 infected patients, ART is not routinely available in developing countries.

SUMMARY OF THE INVENTION

[0005] In certain embodiments, the invention provides compositions and methods for induction of immune response, for example cross-reactive (broadly) neutralizing (bn) Ab induction. In certain embodiments, the methods use compositions comprising HIV-1 immunogens designed to bind to precursors, and/or UCAs of different HIV-1 bnAbs. In certain embodiments, these are UCAs of V1V2 glycan and V3 glycan antibodies.

[0006] In certain aspects, the invention provides compositions comprising a selection of HIV-1 envelopes and/or nucleic acids encoding these envelopes as described herein for example but not limited to selections as described herein. Without limitations, these selected

combinations comprise envelopes which provide representation of the sequence (genetic) and antigenic diversity of the HIV-1 envelope variants which lead to the induction of V1V2 glycan and V3 glycan antibody lineages.

[0007] In certain embodiments the invention provides a recombinant CON-S HIV-1 envelope comprising substitutions at positions N138 andN141 (HXB2 numbering and Figure 16) so that the envelope lacks glycosylation at these positions, and in some embodiments is glycosylated at N301 (HXB2 numbering and Figure 16) and N332 (HXB2 numbering and Figure 16). In certain embodiments the invention provides a recombinant CON-S HIV-1 envelope comprising substitutions at positions N130, N135, N138, and N141 (HXB2 numbering and Figure 16) so that the envelope lacks glycosylation at these positions, and in some embodiments is glycosylated at N301 (HXB2 numbering and Figure 16) and N332 (HXB2 numbering and Figure 16).

[0008] In certain aspects, the invention provides a recombinant CON-S HIV-1 envelope comprising a V1 region, wherein the envelope lacks glycosylation at position N138 and N141 (HXB2 numbering and Figure 16) and in some embodiments is glycosylated at N301 (HXB2 numbering and Figure 16) and N332 (HXB2 numbering and Figure 16).

[0009] In certain aspects, the invention provides a recombinant CON-S HIV-1 envelope comprising a V1 region, wherein the envelope lacks glycosylation at position N130, N135, N138, and N141 (HXB2 numbering and Figure 16) and in some embodiments is glycosylated at N301 (HXB2 numbering) and N332 (HXB2 numbering and Figure 16). The recombinant CON-S HIV-1 envelope further lacking glycosylation at position N138 and N141 (HXB2 numbering and Figure 16) and in some embodiments is glycosylated at N301 (HXB2 numbering and Figure 16) and N332 (HXB2 numbering and Figure 16).

[0010] In certain embodiments, the CON-S HIV-1 envelope comprises the following substitutions at positions N130D, N135K, N138S, and N141S, and in some embodiments is glycosylated at N301 (HXB2 numbering) and N332 (HXB2 numbering). In certain embodiments, the CON-S HIV-1 envelope comprises the following substitutions at positions N138S and N141S, and in some embodiments is glycosylated at N301 (HXB2 numbering) and N332 (HXB2 numbering). Any other suitable amino acid substitution (X) is contemplated at these positions. In some embodiments, substitutions are made such that the structure and activity of CON-S is not substantially affected other than the glycosylation at the modified positions. In some embodiments, the substitutions could be selected from amino acids naturally occurring at these positions, wherein these positions are not glycosylated.

[0011] In certain aspects, the invention provides a recombinant CON-S HIV-1 envelope comprising a 17aa-long V1 region, wherein the envelope lacks glycosylation at positions N138 and N141 (HXB2 numbering) and in some embodiments is glycosylated at N301 (HXB2 numbering) and N332 (HXB2 numbering).

[0012] In certain aspects, the invention provides a recombinant CON-S HIV-1 envelope comprising a 17aa-long V1 region, wherein the envelope lacks glycosylation at positions N130, N135, N138, and N141 (HXB2 numbering) and in some embodiments is glycosylated at N301 (HXB2 numbering) and N332 (HXB2 numbering).

[0013] In certain embodiments, the recombinant envelope binds to precursors, and/or UCAs of different HIV-1 bnAbs. In certain embodiments, these are UCAs of V1V2 glycan and V3 glycan antibodies.

[0014] In certain embodiments, the envelope comprises all consecutive amino acids after the signal peptide of CON-Schim.6R.DS.SOSIP.664_OPT_N138X_N141X (Table 3).

[0015] In certain embodiments, the envelope comprises all consecutive amino acids after the signal peptide of CON-Schim.6R.DS.SOSIP.664_OPT_N138S_N141S (Table 3, Figure 10).

[0016] In certain embodiments, the envelope comprises all consecutive amino acids after the signal peptide of CON-Schim.6R.DS.SOSIP.664_OPT_N130X_N135X_N138X_N141X_ (Table 3).

[0017] In certain embodiments, the envelope comprises all consecutive amino acids after the signal peptide of CON-Schim.6R.DS.SOSIP.664_OPT_N130D_N135K_N138S_N141S_ (Table 3, Figure 10).

[0018] In certain embodiments, the envelope comprises all consecutive amino acids after the signal peptide of CON-Schim.6R.DS.SOSIP.664_OPT_N130X_N135X_N138X_N141X_ferritin (Table 3).

[0019] In certain embodiments, the envelope comprises all consecutive amino acids after the signal peptide of CON-Schim.6R.DS.SOSIP.664_OPT_N130D_N135K_N138S_N141S_ferritin (Table 3, Figure 10, Figure 11B).

[0020] In certain embodiments, the envelope comprises all consecutive amino acids after the signal peptide of CON-Schim.6R.DS.SOSIP.664_OPT_D1305V1 (Table 1, Figure 13).

[0021] In certain aspects, the invention provides a recombinant CON-S envelope wherein the envelope is a protomer comprised in a stable trimer.

[0022] In certain embodiments, the envelope is any one of the envelopes of Table 1, 2 or 3. A skilled artisan appreciates that when recombinantly expressed the envelope proteins do not

include a signal sequence. The CON-S recombinant of any of the embodiments, wherein the envelope is designed so that it forms a multimer, wherein the multimer is a trimer, or in some embodiment, the multimer comprises several trimers, e.g. but not limited to trimers arrayed in a ferritin nanoparticle.

[0023] In certain embodiments, the recombinant Con-S envelope is Man9GlcNAc-enriched. In some embodiments the envelopes are recombinantly produced under kif treatment.

[0024] In certain embodiments, the envelope comprises additional mutations stabilizing the trimer. In certain embodiments these including but are not limited to SOSIP mutations. In certain embodiments mutations are selected from sets F1-F14, VT1-VT8 mutations described herein, or any combination or subcombination within a set.

[0025] In certain aspects, the invention provides a composition comprising any one the recombinant CON-S envelopes and a carrier.

[0026] In certain aspects, the invention provides a nucleic acid encoding any one of the recombinant envelopes of the invention. In certain aspects, the invention provides a composition comprising a nucleic acid encoding the recombinant proteins of the invention and a carrier. In certain embodiments, the nucleic acid is a modified mRNA.

[0027] In certain aspects, the invention provides composition comprising the recombinant CON-S envelope of the invention, wherein the recombinant CON-S envelope is multimerized and wherein optionally in some embodiments the envelope is comprised in a nanoparticle.

[0028] In certain embodiments, the recombinant CON-S envelope is comprised in a nanoparticle that is a ferritin nanoparticle.

[0029] In certain aspects, the invention provides methods of inducing an immune response in a subject comprising administering an immunogenic composition comprising the recombinant CON-S envelopes of the invention or a composition comprising the inventive recombinant CON-S envelopes. In certain embodiments, the methods comprise administering a series of immunogenic compositions, a non-limiting embodiment shown in Figure 3A.

[0030] In certain embodiments of the methods, the immunogenic composition comprises a first immunogen comprising all consecutive amino acids after the signal peptide of CON-Schim.6R.DS.SOSIP.664_OPT_N130D_N135K_N138S_N141S_ferritin (Table 3, Figure 10, Figure 11B), and wherein the immunogen is optionally Man9GlcNAc-enriched, optionally multimerized in a nanoparticle, wherein the nanoparticle is a ferritin nanoparticle.

[0031] In certain embodiments, the methods further comprise administering a second immunogenic composition comprising a second immunogen comprising all consecutive amino acids after the signal peptide of CON-

Schim.6R.DS.SOSIP.664_OPT_N130D_N135K_N138S_N141S_ (Table 3, Figure 10, Figure 11B), and wherein the immunogen is optionally multimerized in a trimer.

[0032] In certain embodiments, the methods further comprise administering a third immunogenic composition comprising a third immunogen comprising all consecutive amino acids after the signal peptide of CON-Schim.6R.DS.SOSIP.664_OPT_N138S_N141S (Table 3, Figure 10), wherein the immunogen is optionally Man9GlcNAc-enriched and optionally multimerized in a trimer.

[0033] In certain embodiments, the methods further comprise administering a fourth immunogenic composition comprising a fourth immunogen comprising all consecutive amino acids after the signal peptide of CON-Schim.6R.DS.SOSIP.664_OPT_N138S_N141S (Table 3, Figure 10), wherein the immunogen is optionally multimerized in a trimer.

[0034] In certain embodiments, the methods further comprise administering a fifth immunogenic composition comprising a fifth immunogen comprising all consecutive amino acids after the signal peptide of CON-Schim.6R.DS.SOSIP.664_OPT_ (Table 3, Figure 10), wherein the immunogen is optionally Man9GlcNAc-enriched and optionally multimerized in a trimer.

[0035] In certain embodiments, the methods further comprise comprising administering a sixth immunogenic composition comprising a sixth immunogen comprising all consecutive amino acids after the signal peptide of CON-Schim.6R.DS.SOSIP.664_OPT_ (Table 3, Figure 10), wherein the immunogen is optionally multimerized in a trimer.

[0036] In certain embodiments, the methods further comprise administering an adjuvant.

[0037] In certain embodiments of the methods, the composition is administered as a prime and/or a boost. In certain embodiments, the composition comprises nanoparticles.

[0038] In certain aspects, the invention provides a recombinant HIV-1 envelope comprising 17aa V1 region, lacks glycosylation at position N133 and N138 (HXB2 numbering), includes glycosylation at N301 (HXB2 numbering) and N332 (HXB2 numbering), and optionally further comprises the “GDIR” motif.

[0039] In certain embodiments, the recombinant envelope binds to precursors, and/or UCAs of different HIV-1 bnAbs. In certain embodiments, these are UCAs of V1V2 glycan and V3 glycan antibodies. In certain embodiments, the envelope is 19CV3. In certain embodiments, the envelope is not 10.17 DT variant described previously in WO/2018161049.

[0040] In certain embodiments, the envelope is a protomer which could be comprised in a stable trimer. In certain embodiments, the envelope comprises additional mutations stabilizing the trimer. In certain embodiments these including but are not limited to SOSIP

mutations. In certain embodiments, mutations are selected from sets F1-F14, VT1-VT8 mutations described herein, or any combination or subcombination within a set.

[0041] In certain embodiments, recombinant HIV-1 envelopes are shown in Figures 10-16. A nucleic acid encoding any of the recombinant envelopes. In certain aspects, the invention provides a composition comprising any one of the inventive envelopes or nucleic acid sequences encoding the same.

[0042] Provides are compositions comprising a nanoparticle which comprises any one of the envelopes of the invention. The composition of any of the embodiments, wherein the nanoparticle is ferritin self assembling nanoparticle.

[0043] In certain aspects, the invention provides methods of inducing an immune response in a subject comprising administering an immunogenic composition comprising any one of the stabilized envelopes of the invention. In certain embodiments, the composition is administered as a prime and/or a boost. In certain embodiments, the composition comprises nanoparticles.

[0044] In certain aspects, the invention provides compositions comprising a plurality of nanoparticles comprising a plurality of the envelopes/trimers of the invention, and a carrier. In non-limiting embodiments, the envelopes/trimers of the invention are multimeric when comprised in a nano-particle. The nanoparticle size is suitable for delivery and purification. In non-liming embodiments, the nanoparticles are ferritin-based nano-particles.

[0045] In certain aspects, the invention provides a ConS envelope where four glycans are removed: N130, N135, N138 and N141. In some embodiments these glycans are removed by the following amino acid changes N130D, N135K, N138S, and N141S. Any ConS envelope comprising any one of these glycan site modifications, or combination thereof is contemplated.

[0046] In certain aspects the invention provides a ConS envelope where two glycans are removed: N138 and N141. In some embodiments these glycans are removed by the following amino acid changes N138S and N141S. Any ConS envelope comprising these glycan site modifications is contemplated.

[0047] In certain embodiments the inventive envelopes are produced recombinantly in the presence of kifunensine. Recombinant envelopes grown in kif treated cells are Man9GlcNAc2 enriched.

BRIEF DESCRIPTION OF THE DRAWINGS

[0048] The patent or application file contains at least one drawing executed in color. To conform to the requirements for PCT patent applications, many of the figures presented herein are black and white representations of images originally created in color.

[0049] Figures 1A-1B show that Kif treatment improves V3-glycan bnAb neutralization. (A) Structural model of HIV-1 Env depicting bnAb epitopes in different colors. (B) Neutralization of untreated and kif-treated JRFL by the V3-glycan bnAb lineage DH270.

[0050] Figures 2A-2E show that V3-glycan bnAbs and precursor antibodies bind optimized CON-S SOSIP trimers. (A-B) 2D-class average of negative stain EM images of (A) kif-treated SOSIP nanoparticles and (B) free trimers. (C) Kif-treated CON-S nanoparticles can bind to unmutated common ancestors (UCAs) of V1V2 and V3-glycan bnAbs. (D) Kif-treated CON-S nanoparticle binds with a higher magnitude to affinity matured DH270 than DH270_UCA3. (E) Kif treatment and nanoparticle presentation enhances DH270 binding to CON-S.

[0051] Figures 3A-3C show rapid induction of CON-S binding antibodies in immunized macaques. (A) Vaccination Regimen. Sequential vaccine regimen aimed at eliciting V3-glycan antibodies. Immunization of rhesus macaques to determine immunogenicity of CON-S SOSIP nanoparticles. (B) 2D-class average negative stain EM images of HIV-1 Env immunogens and antigenicity profile. (C) Serum IgG binding titers to CON-S and CH848 envelopes and ferritin. Arrows indicate immunization time points.

[0052] Figure 4 shows that vaccinated rhesus macaque serum blocks the binding of V3-glycan bnAb DH270 and glycan bnAb 2G12 to Env. Serum blocking was measured in a competition ELISA. Greater than 20% is considered positive. Blocking arose quicker on Envs where V1 glycans were removed.

[0053] Figure 5 shows vaccination induced N301 glycan-dependent autologous tier 2 neutralization. Serum neutralization of autologous tier 2 viruses arose after three immunizations in all 4 rhesus macaques. A N301A mutation decreased neutralization in 2 of 4 macaques. Kifunensine-dependent serum neutralization of JRFL was also elicited.

[0054] Figure 6 shows one embodiment of Con-S vaccine design.

[0055] Figure 7 shows Preliminary ConS glycans.

[0056] Figure 8 shows CON-S vs BG505 SOSIPs glycans.

[0057] Figures 9A-9B show Disulfide Bond Topology: CON-Schim.6R.DS.SOSIP.664_avi _OPT_N130D_N135K_N138S_N141S (mutant). Figure 9A shows canonical Disulfide Bond Topology. Figure 9B shows alternative Disulfide Bond Topology.

[0058] Figure 10 shows non-limiting embodiments of sequences.

[0059] Figure 11A shows a non-limiting embodiment of a Con-S sequence SOSIP design with a modified V1 loop. In one embodiment, the V1 loop is from a naturally occurring envelope CH848.3.D1305.10.19. Bolded is the position of the V1 loop; some amino acids are shared, i.e. are the same between the ConS V1 loop and 10.19 V1 loop, but the ones that are different are modified in ConS. In some embodiments of the invention, the 10.19 V1 loop can replace the V1 loop in any of the ConS sequences and designs.

[0060] Figure 11B shows a non-limiting embodiment of an HIV-1 envelope comprising a ferritin sequence for multimerization. This sequence comprises as annotated: a cloning site at the beginning of the sequence, indicated by the italicized sequence in this figure, a signal peptide, indicated by the underlined position, and one embodiment of a liker between the envelope sequences and the ferritin protein, indicated by the bolded amino acids.

[0061] Figure 12 shows non-limiting embodiments of sequences.

[0062] Figure 13 shows non-limiting embodiments of sequences.

[0063] Figure 14 shows non-limiting embodiments of sequences.

[0064] Figure 15 shows the sequence the sortase A tagged SOSIP trimer HV1301580_C_SORTA;CH848.3.D1305.10.19_D949V3.DS.SOSIP_C_SORTA. The sequence is of sortase A tagged SOSIP trimer. The Sortase A tag is LPSTGG which is modified from LPSTG because an additional Gly residue helps accelerate the reaction rate.

[0065] Figure 16 shows sequences of CON-S gp160 envelope with four N glycosylation sites (N130, N135, N138 and N141 bolded and underlined).

[0066] Figure 17A-17B shows schematic of an HIV envelope SOSIP ferritin nanoparticles by sortase-A conjugation. Figure 17A. Diagram of one non-limiting embodiment of HIV-1 envelope SOSIP trimer showing the orientation of sortase A linkage to ferritin. The sortase linker in this embodiment is LPSTGG. Figure 17B. A model of an HIV-1 env SOSIP ferritin particle with 8 Env trimers displayed, based on ferritin and SOSIP trimer crystal structures.

[0067] Figure 18A-C Ca2+ flux experiments. These data show that Con-S delta V1 glycans triggers DH270 IA4 but JRFL degly ("Q3") does not. Figure 18A. Experiment Name: CaF020; Protein: CON-Sgp140CFI.avi/293F/Trimer; Concentration Assayed: 0.100nMolar of protein tetramer; Data presented as % of max anti-IgM Fab(2) 50ug/mL. Figure 18B. Experiment Name: CaF020; Protein: JRFLgp140CF.avi V1 3Q/293F/Trimer; Concentration Assayed: 0.100nMolar of protein tetramer; Data presented as % of max anti-IgM Fab(2) 50ug/mL. Figure 18C. Experiment Name: CaF020; Protein: CON-

Sgp140CFI.avi/293F/Trimer; Concentration Assayed: 0.100nMolar of protein tetramer; Data presented as % of max anti-IgM Fab(2) 50ug/mL.

[0068] Figures 19A-19C show expression of sequential CON-S stabilized SOSIP trimers with V1 glycans present (A) or removed, (B) two glycans N138S and N141S are removed and (C) four glycans N130D, N135K, N138S, and N141S are removed. Top panel shows size exclusion chromatography and bottom panel shows negative stain EM. Expression of stabilized CON-S SOSIP gp140 trimers with serially deleted V1 glycosylation sites. Size exclusion chromatography shows most of the protein purified with PGT145 affinity chromatography is trimeric Env. Trimeric envelope was visualized by negative stain electron microscopy and 2-dimensional class averaging.

[0069] Figure 20 shows Glycosylation profile of the stabilized CON-S SOSIP. The data shows that CON-S SOSIPs are glycosylated with only high mannose at the N332 glycan bnAb epitope. The same glycosylation profile was obtained when the four V1 glycans were removed. The glycosylation profile of CON-S SOSIP with (CON-S gp140 chim. 6R.SOSIP.664.avi) and without V1 glycans at N130, N135, N138, N141 (CON-S gp140 chim. 6R.SOSIP.664.avi_opt mutant). Mass spectrometry shows that the CON-S SOSIP gp140 with N130D, N135K, N138S, and N141S mutations lacks any glycans at N130, N135, N138, N141. The percentage of high mannose at neutralizing antibody contact sites 295, 301, 332, 156, and 611 does not change with removal of the glycans.

[0070] Figures 21A-21B show CON-S Ferritin nanoparticles by SOSIP-ferritin fusion proteins. (A) shows SOSIP-Ferritin, (B) shows 2D class average. Fusion of CON-S SOSIP N130D, N135K, N138S, and N141S to *H. pylori* ferritin to create nanoparticles. Negative stain electron microscopy shows nanoparticle formation of CON-S SOSIP

[0071] Figure 22 shows Glycan-modified CON-S binds to V1V2-glycan and V3-glycan bnAb UCAs. The data show V1V2 glycan and V3-glycan bnAb UCA antigenicity. CON-S SOSIP N130D, N135K, N138S, and N141S ferritin nanoparticle is antigenic for V3-glycan bnAb unmutated common ancestors (UCAs) antibodies and V1V2-glycan bnAb precursor CH103 UCA antibody. V3-glycan bnAbs are BF520, BG18, and DH270. DH272 is a V3-glycan antibody that neutralizes only autologous viruses. CH01 is a V1V2-glycan bnAb. Binding was measured by biolayer interferometry with the nanoparticle in solution and each antibody immobilized on a sensor tip. The red vertical line indicates the end of the association phase.

[0072] Figures 23A-23B show that Man9GlcNAc2 enrichment on CON-S ferritin nanoparticles augments V3-glycan bnAb binding. (A) shows kif treated CON-S SOSIP

ferritin nanoparticle. (B) shows DH270 V3 glycan antibody binding to CON-S. Kifunensine treatment enhances V3-glycan bnAb binding to CON-S SOSIP N130D, N135K, N138S, and N141S ferritin nanoparticle. Binding of V3-glycan bnAb DH270 to CON-S SOSIP N130D, N135K, N138S, and N141S was compared to binding to CON-S SOSIP N130D, N135K, N138S, and N141S ferritin nanoparticle. To enrich the glycans on the CON-S SOSIP N130D, N135K, N138S, and N141S ferritin nanoparticle for Man9GlcNAc2 or Man8GlcNAc2 the protein was expressed in cells treated with the glycosidase inhibitor kifunensine.

[0073] Figures 24A-24B show that a single ferritin nanoparticle of CON-S SOSIP lacking four glycans immunization elicited gradually increasing serum IgG over 6 weeks. Numbers on the x-axis show the study week. Arrows indicate immunization.

[0074] Figure 25 shows serum IgG binding to CON-S after trimer boost (II in Figure 3A) was blocked by V1 glycans on the CON-S SOSIP.

[0075] Figures 26A-26B show that NHP serum antibodies block V3-glycan mAb DH270 (A) and glycan mAb 2G12 (B) binding to CON-S SOSIP. Numbers on the x-axis show the study week. Arrows indicate immunization. Plasma antibodies from CON-S SOSIP vaccinated macaques blocks V3-glycan bnAb DH270 and gp120 glycan bnAb 2G12 binding to CON-S SOSIP gp140.

[0076] Figure 27 shows that CON-S SOSIP vaccination induced autologous tier 2 neutralization. Numbers on the x-axis show the study week. Arrows indicate immunization.

[0077] Figure 28 shows that autologous tier 2 neutralization is not dependent on the N362 glycan hole. The glycan shield of CON-S is intact except at N362.

[0078] Figures 29A-29B show that two of the NHPs in the study in Example 2 generated N301-dependent autologous tier 2 neutralizing antibodies (compare 29A and 29B). The figure shows CON-S neutralization (heterogenous glycans). Numbers on the x-axis show the study week. Arrows indicate immunization. Immunization induces autologous tier 2 neutralizing antibodies in two macaques that target the N301 glycan in the V3-glycan epitope on Env. Neutralization titer was measured in the TZM-bl assay as serum dilution that inhibits 50% of virus replication (ID50). Solid lines are the neutralization titers for wildtype CON-S. Dashed lines are neutralization titers for CON-S with the asparagine301 (N301) mutated to alanine to disrupt the glycosylation sequence. Arrows on the x-axis indicate immunization timepoints.

[0079] Figures 30A-30F show BLI data for 442EML(b) CON_Schim.6R.DS.SOSIP.664_OPT_N130D_N135K_N138S_ (kif treated envelope)--(50ug/mL). Vertical dotted lines indicate response at end of association. Panels show

biolayer interferometry binding of a panel of HIV-1 antibodies to the CON-S envelope nanoparticle. The data shows that there is trimeric envelope present on the nanoparticle and that the envelope presents four bnAb epitopes. The cells producing the nanoparticle are treated with kifunensine to enrich for Man9GlcNac2 glycans during protein synthesis. The CON-S protein binds to trimer-specific antibody PGT145. It weakly binds to trimer-specific antibody PGT151 because the Env glycosylation has been restricted to Man9GlcNAc2 and PGT151 requires complex glycans. The envelope is not antigenic for the inferred precursor of the CH106 lineage (CH103 UCA) but can bind to somatically mutated broadly neutralizing antibodies against the CD4 binding site such as VRC01 and CH106. The envelope is not antigenic for antibodies that recognize the CD4-induced conformation of Env (A32, 17B, and CH58). Antibodies against the conformational V3 glycan bnAb epitope bind to the envelope showing this envelope has a well-folded V3 loop and the prototypical HIV-1 envelope high mannose patch is present on the envelope. The binding of 19B and F39F shows that not all of the envelope is the closed conformation, which has the V3 loop inaccessible to 19B and F39F.

[0080] Figure 31 shows a summary CON-Schim.6R.DS.SOSIP.664_OPT_N130D_N135K_N138S_N141S_ferritin/5uM-Kif/293F 442EML Binding Results. Panels show biolayer interferometry binding of a panel of HIV-1 antibodies to the CON-S envelope nanoparticle. The data shows that there is trimeric envelope present on the nanoparticle and that the envelope presents four bnAb epitopes. The CON-S protein binds to trimer-specific antibodies PGT145 and PGT151. The envelope is not antigenic for the inferred precursor of the CH106 lineage (CH103 UCA) but can bind to somatically mutated broadly neutralizing antibodies against the CD4 binding site such as VRC01 and CH106. The envelope is not antigenic for antibodies that recognize the CD4-induced conformation of Env (A32, 17B, and CH58). Antibodies against the conformational V3 glycan bnAb epitope bind to the envelope showing this envelope has a well-folded V3 loop and the protypical HIV-1 envelope high mannose patch is present on the envelope. The binding of 19B and F39F shows that not all of the envelope is the closed conformation, which has the V3 loop inaccessible to 19B and F39F.

[0081] Figures 32A-32F BLI data for 455EML CONSchim.6R.DS.SOSIP.664_OPT_N130D_N135K_N138S_N141 (kif untreated)--(50ug/mL). Vertical dotted lines indicate response at end of association.

[0082] Figure 33 shows a summary CON-Schim.6R.DS.SOSIP.664_OPT_N130D_N135K_N138S_N141S_ferritin/293F

[0083] 455EML 17-July-2018 Binding Results.

[0084] Figure 34 shows a summary of antigenic profile of CON-S envelopes with various V1 glycosylation sites removed.

DETAILED DESCRIPTION OF THE INVENTION

[0085] The development of a safe, highly efficacious prophylactic HIV-1 vaccine is of paramount importance for the control and prevention of HIV-1 infection. A major goal of HIV-1 vaccine development is the induction of broadly neutralizing antibodies (bnAbs) (Immunol. Rev. 254: 225-244, 2013). BnAbs are protective in rhesus macaques against SHIV challenge, but as yet, are not induced by current vaccines.

[0086] The invention provides methods of using these pan bnAb envelope immunogens.

[0087] In certain aspect, the invention provides compositions for immunizations to induce lineages of broad neutralizing antibodies. In certain embodiments, there is some variance in the immunization regimen; in some embodiments, the selection of HIV-1 envelopes may be grouped in various combinations of primes and boosts, either as nucleic acids, proteins, or combinations thereof. In certain embodiments, the compositions are pharmaceutical compositions which are immunogenic. In certain embodiments, the compositions comprise amounts of envelopes which are therapeutic and/or immunogenic.

[0088] In one aspect the invention provides a composition for a prime boost immunization regimen comprising any one of the envelopes described herein, or any combination thereof wherein the envelope is a prime or boost immunogen. In certain embodiments, the composition for a prime boost immunization regimen comprises one or more envelopes described herein.

[0089] In certain embodiments, the compositions contemplate nucleic acid, as DNA and/or RNA, or proteins immunogens either alone or in any combination. In certain embodiments, the methods contemplate genetic, as DNA and/or RNA, immunization either alone or in combination with envelope protein(s).

[0090] mRNA

[0091] In some embodiments, the antigens are nucleic acids, including but not limited to mRNAs which could be modified and/or unmodified. See US Pub 20180028645A1, US Pub 20170369532, US Pub 20090286852, US Pub 20130111615, US Pub 20130197068, US Pub 20130261172, US Pub 20150038558, US Pub 20160032316, US Pub 20170043037, US Pub

20170327842, each content is incorporated by reference in its entirety. mRNAs delivered in LNP formulations have advantages over non-LNPs formulations. See US Pub 20180028645A1.

[0092] In certain embodiments, the nucleic acid encoding an envelope is operably linked to a promoter inserted an expression vector. In certain aspects, the compositions comprise a suitable carrier. In certain aspects, the compositions comprise a suitable adjuvant.

[0093] In certain embodiments, the induced immune response includes induction of antibodies, including but not limited to autologous and/or cross-reactive (broadly) neutralizing antibodies against HIV-1 envelope. Various assays that analyze whether an immunogenic composition induces an immune response, and the type of antibodies induced are known in the art and are also described herein.

[0094] In certain aspects, the invention provides an expression vector comprising any of the nucleic acid sequences of the invention, wherein the nucleic acid is operably linked to a promoter. In certain aspects, the invention provides an expression vector comprising a nucleic acid sequence encoding any of the polypeptides of the invention, wherein the nucleic acid is operably linked to a promoter. In certain embodiments, the nucleic acids are codon optimized for expression in a mammalian cell, in vivo or in vitro. In certain aspects, the invention provides nucleic acids comprising any one of the nucleic acid sequences of invention. In certain aspects, the invention provides nucleic acids consisting essentially of any one of the nucleic acid sequences of invention. In certain aspects, the invention provides nucleic acids consisting of any one of the nucleic acid sequences of invention. In certain embodiments the nucleic acid of the invention, is operably linked to a promoter and is inserted in an expression vector. In certain aspects, the invention provides an immunogenic composition comprising the expression vector.

[0095] In certain aspects, the invention provides a composition comprising at least one of the nucleic acid sequences of the invention. In certain aspects, the invention provides a composition comprising any one of the nucleic acid sequences of invention. In certain aspects, the invention provides a composition comprising at least one nucleic acid sequence encoding any one of the polypeptides of the invention.

[0096] The envelope used in the compositions and methods of the invention can be a gp160, gp150, gp145, gp140, gp120, gp41, N-terminal deletion variants as described herein, cleavage resistant variants as described herein, or codon optimized sequences thereof. In certain embodiments, the composition comprises envelopes as trimers. In certain embodiments, envelope proteins are multimerized, for example, trimers are attached to a

particle such that multiple copies of the trimer are attached and the multimerized envelope is prepared and formulated for immunization in a human. In certain embodiments, the compositions comprise envelopes, including but not limited to trimers as particulate, high-density array on liposomes or other particles, for example but not limited to nanoparticles. In some embodiments, the trimers are in a well ordered, near native like or closed conformation. In some embodiments, the trimer compositions comprise a homogenous mix of native like trimers. In some embodiments, the trimer compositions comprise at least 85%, 90%, or 95% native like trimers.

[0097] In certain embodiments, the envelope is any of the forms of HIV-1 envelope. In certain embodiments, the envelope is gp120, gp140, gp145 (i.e. with a transmembrane), or gp150. In certain embodiments, gp140 designed to form a stable trimer. In certain embodiments envelope protomers from a trimer which is not a SOSIP timer. In certain embodiment, the trimer is a SOSIP based trimer wherein each protomer comprises additional modifications. In certain embodiments, envelope trimers are recombinantly produced. In certain embodiments, envelope trimers are purified from cellular recombinant fractions by antibody binding and reconstituted in lipid comprising formulations. See for example WO2015/127108 titled "Trimeric HIV-1 envelopes and uses thereof" and WO/2017151801which content is herein incorporated by reference in its entirety. In certain embodiments, the envelopes of the invention are engineered and comprise non-naturally occurring modifications.

[0098] In certain embodiments, the envelope is in a liposome. In certain embodiments, the envelope comprises a transmembrane domain with a cytoplasmic tail embedded in a liposome. In certain embodiments, the nucleic acid comprises a nucleic acid sequence, which encodes a gp120, gp140, gp145, gp150, or gp160.

[0099] In certain embodiments, where the nucleic acids are operably linked to a promoter and inserted in a vector, the vector is any suitable vector. Non-limiting examples include, VSV, replicating rAdenovirus type 4, MVA, Chimp adenovirus vectors, pox vectors, and the like. In certain embodiments, the nucleic acids are administered in NanoTaxi block polymer nanospheres. In certain embodiments, the composition and methods comprise an adjuvant. Non-limiting examples include, 3M052, AS01 B, AS01 E, gla/SE, alum, Poly I poly C (poly IC), polyIC/long chain (LC) TLR agonists, TLR7/8 and 9 agonists, or a combination of TLR7/8 and TLR9 agonists (see Moody et al. (2014) J. Virol. March 2014 vol. 88 no. 6 3329-3339), or any other adjuvant. Non-limiting examples of TLR7/8 agonist include TLR7/8 ligands, Gardiquimod, Imiquimod and R848 (resiquimod). A non-limiting embodiment of a

combination of TLR7/8 and TLR9 agonist comprises R848 and oCpG in STS (see Moody et al. (2014) J. Virol. March 2014 vol. 88 no. 6 3329-3339).

[0100] In certain aspects the invention provides a cell comprising a nucleic acid encoding any one of the envelopes of the invention suitable for recombinant expression. In certain aspects, the invention provides a clonally derived population of cells encoding any one of the envelopes of the invention suitable for recombinant expression. In certain aspects, the invention provides a sable pool of cells encoding any one of the envelopes of the invention suitable for recombinant expression.

[0101] In certain aspects, the invention provides a recombinant HIV-1 envelope polypeptide as described here, wherein the polypeptide is a non-naturally occurring protomer designed to form an envelope trimer. The invention also provides nucleic acids encoding these recombinant polypeptides. Non-limiting examples of amino acids and nucleic acid of such protomers are referenced in Tables 1-3, and Figures 10-16.

[0102] In certain aspects the invention provides a recombinant trimer comprising three identical protomers of an envelope. In certain aspects, the invention provides an immunogenic composition comprising the recombinant trimer and a carrier, wherein the trimer comprises three identical protomers of an HIV-1 envelope as described herein. In certain aspects, the invention provides an immunogenic composition comprising nucleic acid encoding these recombinant HIV-1 envelope and a carrier.

[0103] Sequences/Clones

[0104] Described herein are nucleic and amino acids sequences of HIV-1 envelopes. The sequences for use as immunogens are in any suitable form. In certain embodiments, the described HIV-1 envelope sequences are gp160s. In certain embodiments, the described HIV-1 envelope sequences are gp120s. Other sequences, for example but not limited to stable SOSIP trimer designs, gp145s, gp140s, both cleaved and uncleaved, gp140 Envs with the deletion of the cleavage (C) site, fusion (F) and immunodominant (I) region in gp41--named as gp140ΔCFI (gp140CFI), gp140 Envs with the deletion of only the cleavage (C) site and fusion (F) domain -- named as gp140ΔCF (gp140CF), gp140 Envs with the deletion of only the cleavage (C)—named gp140ΔC (gp140C) (See e.g. Liao et al. **Virology 2006, 353, 268**-282), gp150s, gp41s, which are readily derived from the nucleic acid and amino acid gp160 sequences. In certain embodiments the nucleic acid sequences are codon optimized for optimal expression in a host cell, for example a mammalian cell, a rBCG cell or any other suitable expression system.

[0105] An HIV-1 envelope has various structurally defined fragments/forms: gp160; gp140---including cleaved gp140 and uncleaved gp140 (gp140C), gp140CF, or gp140CFI; gp120 and gp41. A skilled artisan appreciates that these fragments/forms are defined not necessarily by their crystal structure, but by their design and bounds within the full length of the gp160 envelope. While the specific consecutive amino acid sequences of envelopes from different strains are different, the bounds and design of these forms are well known and characterized in the art.

[0106] For example, it is well known in the art that during its transport to the cell surface, the gp160 polypeptide is processed and proteolytically cleaved to gp120 and gp41 proteins. Cleavages of gp160 to gp120 and gp41 occurs at a conserved cleavage site "REKR." *See* Chakrabarti et al. Journal of Virology vol. 76, pp. 5357-5368 (2002) see for example Figure 1, and Second paragraph in the Introduction on p. 5357; Binley et al. Journal of Virology vol. 76, pp. 2606-2616 (2002) for example at Abstract; Gao et al. Journal of Virology vol. 79, pp. 1154-1163 (2005); Liao et al. Virology vol. 353(2): 268–282 (2006).

[0107] The role of the furin cleavage site was well understood both in terms of improving cleave efficiency, *see* Binley et al. supra, and eliminating cleavage, s*ee* Bosch and Pawlita, Virology 64 (5):2337-2344 (1990); Guo et al. Virology 174: 217-224 (1990); McCune et al. Cell 53:55-67 (1988); Liao et al. J Virol. Apr;87(8):4185-201 (2013).

[0108] Likewise, the design ofgp140 envelope forms is also well known in the art, along with the various specific changes which give rise to the gp140C (uncleaved envelope), gp140CF and gp140CFI forms. Envelope gp140 forms are designed by introducing a stop codon within the gp41 sequence. *See* Chakrabarti et al. at Figure 1.

[0109] Envelope gp140C refers to a gp140 HIV-1 envelope design with a functional deletion of the cleavage (C) site, so that the gp140 envelope is not cleaved at the furin cleavage site. The specification describes cleaved and uncleaved forms, and various furin cleavage site modifications that prevent envelope cleavage are known in the art. In some embodiments of the gp140C form, two of the R residues in and near the furin cleavage site are changed to E, e.g., **R**RVVEREK**R is** changed to **E**RVVEREK**E,** and is one example of an uncleaved gp140 form. Another example is the gp140C form which has the REKR site changed to SEKS. *See* supra for references.

[0110] Envelope gp140CF refers to a gp140 HIV-1 envelope design with a deletion of the cleavage (C) site and fusion (F) region. Envelope gp140CFI refers to a gp140 HIV-1 envelope design with a deletion of the cleavage (C) site, fusion (F) and immunodominant (I) region in gp41. *See* Chakrabarti et al. Journal of Virology vol. 76, pp. 5357-5368 (2002) see

for example Figure 1, and Second paragraph in the Introduction on p. 5357; Binley et al. Journal of Virology vol. 76, pp. 2606-2616 (2002) for example at Abstract; Gao et al. Journal of Virology vol. 79, pp. 1154-1163 (2005); Liao et al. Virology vol. 353(2): 268–282 (2006).

[0111] In certain embodiments, the envelope design in accordance with the present invention involves deletion of residues (e.g., 5-11, 5, 6, 7, 8, 9, 10, or 11 amino acids) at the N-terminus. For delta N-terminal design, amino acid residues ranging from 4 residues or even fewer to 14 residues or even more are deleted. These residues are between the maturation (signal peptide, which can be readily determined by a skilled artisan) and "VPVXXXX…". In case of ConS Env as an example, amino acids (italicized and underlined in the below sequence) were deleted between the signal peptide and the VPVXXX..: MGSLQPLATLYLLGMLVASVLAAENLWVTVYYGVPVWKEANTT…(rest of envelope sequence is indicated as "…"). In certain embodiments, the invention relates generally to an immunogen, gp160, gp120 or gp140, without an N-terminal Herpes Simplex gD tag substituted for amino acids of the N-terminus of gp120, with an HIV leader sequence (or other leader sequence), and without the original about 4 to about 25, for example 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 amino acids of the N-terminus of the envelope (e.g. gp120). See WO2013/006688, e.g. at pages 10-12, the contents of which publication is hereby incorporated by reference in its entirety.

[0112] The general strategy of deletion of N-terminal amino acids of envelopes results in proteins, for example gpl20s, expressed in mammalian cells that are primarily monomeric, as opposed to dimeric, and, therefore, solves the production and scalability problem of commercial gp120 Env vaccine production. In other embodiments, the amino acid deletions at the N-terminus result in increased immunogenicity of the envelopes.

[0113] In certain embodiments, the invention provides envelope sequences, amino acid sequences and the corresponding nucleic acids, and in which the V3 loop is substituted with the following V3 loop sequence TRPNNNTRKSIRIGPGQTFY ATGDIIGNIRQAH. This substitution of the V3 loop reduced product cleavage and improves protein yield during recombinant protein production in CHO cells.

[0114] In certain aspects, the invention provides composition and methods which use a selection of Envs, as gp120s, gp 140s cleaved and uncleaved, gp145s, gp150s and gp160s, stabilized and/or multimerized trimers, as proteins, DNAs, RNAs, or any combination thereof, administered as primes and boosts to elicit immune response. Envs as proteins would be co-administered with nucleic acid vectors containing Envs to amplify antibody induction. In certain embodiments, the compositions and methods include any immunogenic

HIV-1 sequences to give the best coverage for T cell help and cytotoxic T cell induction. In certain embodiments, the compositions and methods include mosaic and/or consensus HIV-1 genes to give the best coverage for T cell help and cytotoxic T cell induction. In certain embodiments, the compositions and methods include mosaic group M and/or consensus genes to give the best coverage for T cell help and cytotoxic T cell induction. In some embodiments, the mosaic genes are any suitable gene from the HIV-1 genome. In some embodiments, the mosaic genes are Env genes, Gag genes, Pol genes, Nef genes, or any combination thereof. See e.g. US Patent No. 7951377. In some embodiments the mosaic genes are bivalent mosaics. In some embodiments the mosaic genes are trivalent. In some embodiments, the mosaic genes are administered in a suitable vector with each immunization with Env gene inserts in a suitable vector and/or as a protein. In some embodiments, the mosaic genes, for example as bivalent mosaic Gag group M consensus genes, are administered in a suitable vector, for example but not limited to HSV2, would be administered with each immunization with Env gene inserts in a suitable vector, for example but not limited to HSV-2.

[0115] In certain aspects the invention provides compositions and methods of Env genetic immunization either alone or with Env proteins to recreate the swarms of evolved viruses that have led to bnAb induction. Nucleotide-based vaccines offer a flexible vector format to immunize against virtually any protein antigen. Currently, two types of genetic vaccination are available for testing—DNAs and mRNAs.

[0116] In certain aspects the invention contemplates using immunogenic compositions wherein immunogens are delivered as DNA. See Graham BS, Enama ME, Nason MC, Gordon IJ, Peel SA, et al. (2013) DNA Vaccine Delivered by a Needle-Free Injection Device Improves Potency of Priming for Antibody and CD8+ T-Cell Responses after rAd5 Boost in a Randomized Clinical Trial. PLoS ONE 8(4): e59340, page 9. Various technologies for delivery of nucleic acids, as DNA and/or RNA, so as to elicit immune response, both T-cell and humoral responses, are known in the art and are under developments. In certain embodiments, DNA can be delivered as naked DNA. In certain embodiments, DNA is formulated for delivery by a gene gun. In certain embodiments, DNA is administered by electroporation, or by a needle-free injection technologies, for example but not limited to Biojector® device. In certain embodiments, the DNA is inserted in vectors. The DNA is delivered using a suitable vector for expression in mammalian cells. In certain embodiments the nucleic acids encoding the envelopes are optimized for expression. In certain embodiments DNA is optimized, e.g. codon optimized, for expression. In certain

embodiments, the nucleic acids are optimized for expression in vectors and/or in mammalian cells. In non-limiting embodiments these are bacterially derived vectors, adenovirus based vectors, rAdenovirus (e.g. Barouch DH, et al. Nature Med. 16: 319-23, 2010), recombinant mycobacteria (e.g. rBCG or M smegmatis) (Yu, JS et al. Clinical Vaccine Immunol. 14: 886-093,2007; ibid 13: 1204-11,2006), and recombinant vaccinia type of vectors (Santra S. Nature Med. 16: 324-8, 2010), for example but not limited to ALVAC, replicating (Kibler KV et al., PLoS One 6: e25674, 2011 nov 9.) and non-replicating (Perreau M et al. J. virology 85: 9854-62, 2011) NYVAC, modified vaccinia Ankara (MVA)), adeno-associated virus, Venezuelan equine encephalitis (VEE) replicons, Herpes Simplex Virus vectors, and other suitable vectors.

[0117] In certain aspects, the invention contemplates using immunogenic compositions wherein immunogens are delivered as DNA or RNA in suitable formulations. Various technologies which contemplate using DNA or RNA, or may use complexes of nucleic acid molecules and other entities to be used in immunization. In certain embodiments, DNA or RNA is administered as nanoparticles consisting of low dose antigen-encoding DNA formulated with a block copolymer (amphiphilic block copolymer 704). See Cany et al., Journal of Hepatology 2011 vol. 54 j 115–121; Arnaoty et al., Chapter 17 in Yves Bigot (ed.), Mobile Genetic Elements: Protocols and Genomic Applications, Methods in Molecular Biology, vol. 859, pp293-305 (2012); Arnaoty et al. (2013) Mol Genet Genomics. 2013 Aug;288(7-8):347-63. Nanocarrier technologies called Nanotaxi® for immunogenic macromolecules (DNA, RNA, Protein) delivery are under development. See for example technologies developed by incellart.

[0118] mRNA

[0119] In some embodiments the antigens are nucleic acids, including but not limited to mRNAs which could be modified and/or unmodified. See US Pub 20180028645A1, US Pub 20170369532, US Pub 20090286852, US Pub 20130111615, US Pub 20130197068, US Pub 20130261172, US Pub 20150038558, US Pub 20160032316, US Pub 20170043037, US Pub 20170327842, each content is incorporated by reference in its entirety. mRNAs delivered in LNP formulations have advantages over non-LNPs formulations. See US Pub 20180028645A1.

[0120] In certain aspects the invention contemplates using immunogenic compositions wherein immunogens are delivered as recombinant proteins. Various methods for production and purification of recombinant proteins, including trimers such as but not limited to SOSIP

based trimers, suitable for use in immunization are known in the art. In certain embodiments recombinant proteins are produced in CHO cells.

[0121] It is readily understood that the envelope glycoproteins referenced in various examples and figures comprise a signal/leader sequence. It is well known in the art that HIV-1 envelope glycoprotein is a secretory protein with a signal or leader peptide sequence that is removed during processing and recombinant expression (without removal of the signal peptide, the protein is not secreted). *See* for example Li et al. Control of expression, glycosylation, and secretion of HIV-1 gp120 by homologous and heterologous signal sequences. Virology 204(1):266-78 (1994) ("Li et al. 1994"), at first paragraph, and Li et al. Effects of inefficient cleavage of the signal sequence of HIV-1 gpl20 on its association with calnexin, folding, and intracellular transport. PNAS 93:9606-9611 (1996) ("Li et al. 1996"), at 9609. Any suitable signal sequence could be used. In some embodiments the leader sequence is the endogenous leader sequence. Most of the gp120 and gp160 amino acid sequences include the endogenous leader sequence. In other non-limiting examples the leader sequence is human Tissue Plasminogen Activator (TPA) sequence, human CD5 leader sequence (e.g. MPMGSLQPLATLYLLGMLVASVLA*)*. Most of the chimeric designs include CD5 leader sequence. A skilled artisan appreciates that when used as immunogens, and for example when recombinantly produced, the amino acid sequences of these proteins do not comprise the leader peptide sequences.

[0122] The immunogenic envelopes can also be administered as a protein prime and/or boost alone or in combination with a variety of nucleic acid envelope primes (e.g., HIV -1 Envs delivered as DNA expressed in viral or bacterial vectors).

[0123] Dosing of proteins and nucleic acids can be readily determined by a skilled artisan. A single dose of nucleic acid can range from a few nanograms (ng) to a few micrograms (μg) or milligram of a single immunogenic nucleic acid. Recombinant protein dose can range from a few μg micrograms to a few hundred micrograms, or milligrams of a single immunogenic polypeptide.

[0124] Administration: The compositions can be formulated with appropriate carriers using known techniques to yield compositions suitable for various routes of administration. In certain embodiments the compositions are delivered via intramascular (IM), via subcutaneous, via intravenous, via nasal, via mucosal routes, or any other suitable route of immunization.

[0125] The compositions can be formulated with appropriate carriers and adjuvants using techniques to yield compositions suitable for immunization. The compositions can include

an adjuvant, such as, for example but not limited to, alum, poly IC, MF-59 or other squalene-based adjuvant, ASOIB, or other liposomal based adjuvant suitable for protein or nucleic acid immunization. In certain embodiments, the adjuvant is GSK AS01E adjuvant containing MPL and QS21. This adjuvant has been shown by GSK to be as potent as the similar adjuvant AS01B but to be less reactogenic using HBsAg as vaccine antigen [Leroux-Roels et al., IABS Conference, April 2013]. In certain embodiments, TLR agonists are used as adjuvants. In certain embodiments, the adjuvant is 3M052. In other embodiment, adjuvants which break immune tolerance are included in the immunogenic compositions.

[0126] In certain embodiments, the compositions and methods comprise any suitable agent or immune modulation, which could modulate mechanisms of host immune tolerance and release of the induced antibodies. In non-limiting embodiments modulation includes PD-1 blockade; T regulatory cell depletion; CD40L hyperstimulation; soluble antigen administration, wherein the soluble antigen is designed such that the soluble agent eliminates B cells targeting dominant epitopes, or a combination thereof. In certain embodiments, an immunomodulatory agent is administered in at time and in an amount sufficient for transient modulation of the subject's immune response so as to induce an immune response which comprises broad neutralizing antibodies against HIV-1 envelope. Non-limiting examples of such agents is any one of the agents described herein: e.g. chloroquine (CQ), PTP1B Inhibitor - CAS 765317-72-4 - Calbiochem or MSI 1436 clodronate or any other bisphosphonate; a Foxo1 inhibitor, e.g. 344355 | Foxo1 Inhibitor, AS1842856 - Calbiochem; Gleevac, anti-CD25 antibody, anti-CCR4 Ab, an agent which binds to a B cell receptor for a dominant HIV-1 envelope epitope, or any combination thereof. In non-limiting embodiments, the modulation includes administering an anti-CTLA4 antibody, OX-40 agonists, or a combination thereof. Non-limiting examples are of CTLA-1 antibody are ipilimumab and tremelimumab. In certain embodiments, the methods comprise administering a second immunomodulatory agent, wherein the second and first immunomodulatory agents are different.

[0127] **Multimeric Envelopes**

[0128] Presentation of antigens as particulates reduces the B cell receptor affinity necessary for signal transduction and expansion (See Baptista et al. EMBO J. 2000 Feb 15; 19(4): 513–520). Displaying multiple copies of the antigen on a particle provides an avidity effect that can overcome the low affinity between the antigen and B cell receptor. The initial B cell receptor specific for pathogens can be low affinity, which precludes vaccines from being able to stimulate and expand B cells of interest. In particular, very few naïve B cells from which

HIV-1 broadly neutralizing antibodies arise can bind to soluble HIV-1 Envelope. Provided are envelopes, including but not limited to trimers as particulate, high-density array on liposomes or other particles, for example but not limited to nanoparticles. See e.g. He et al. Nature Communications 7, Article number: 12041 (2016), doi:10.1038/ncomms12041; Bamrungsap et al. Nanomedicine, 2012, 7 (8), 1253-1271.

[0129] To improve the interaction between the naïve B cell receptor and immunogens, envelope designed can be created to wherein the envelope is presented on particles, e.g. but not limited to nanoparticle. In some embodiments, the HIV-1 Envelope trimer could be fused to ferritin. Ferritin protein self assembles into a small nanoparticle with three-fold axis of symmetry. At these axes the envelope protein is fused. Therefore, the assembly of the three-fold axis also clusters three HIV-1 envelope protomers together to form an envelope trimer. Each ferritin particle has 8 axes which equates to 8 trimers being displayed per particle. See e.g. Sliepen et al. Retrovirology201512:82, DOI: 10.1186/s12977-015-0210-4; See also Figure 24H-J.

[0130] Another approach to multimerize expression constructs uses staphylococcus Sortase A transpeptidase ligation to conjugate inventive envelope trimers to cholesterol. The trimers can then be embedded into liposomes via the conjugated cholesterol. To conjugate the trimer to cholesterol either a C-terminal LPXTG(X1) tag, wherein X1 could be a Glycine (G), or a N-terminal pentaglycine repeat tag is added to the envelope trimer gene. Cholesterol is also synthesized with these two tags. Sortase A is then used to covalently bond the tagged envelope to the cholesterol. The sortase A-tagged trimer protein can also be used to conjugate the trimer to other peptides, proteins, or fluorescent labels. In non-limiting embodiments, the sortase A tagged trimers are conjugated to ferritin to form nanoparticles.

[0131] Multimerization of SOSIP trimers. Previous strategies for multimerizing SOSIP Env trimers have been successful, but limited in their multivalency. The B cell receptor recognizes and internalizes low-affinity antigens at a greater magnitude when the low-affinity antigen is presented as a multimeric particle as opposed to monomeric protein in solution (Batista FD, Neuberger MS. B cells extract and present immobilized antigen: implications for affinity discrimination. EMBO J. 2000;19(4):513-20). In vivo, the multimerization of HIV-1 Env has improved neutralizing antibody titers in rabbits (Ingale J, Stano A, Guenaga J, Sharma SK, Nemazee D, Zwick MB, et al. High-Density Array of Well-Ordered HIV-1 Spikes on Synthetic Liposomal Nanoparticles Efficiently Activate B Cells. Cell Rep. 2016;15(9):1986-99.) and monkeys (Martinez-Murillo P, Tran K, Guenaga J, Lindgren G, Adori M, Feng Y, et al. Particulate Array of Well-Ordered HIV Clade C Env Trimers Elicits

Neutralizing Antibodies that Display a Unique V2 Cap Approach. Immunity. 2017;46(5):804-17 e7).

[0132] In certain embodiments, there are methods for expressing and purifying the Env trimers as multimers as ferritin nanoparticles. Purification of SOSIP gp140-ferritin fusion proteins can be complicated by the presence of well-folded and poorly folded trimeric Env on the same nanoparticle, so we developed a two-step ferritin assembly process where we first purified well-folded SOSIP gp140 trimers and separately purified ferritin nanoparticles. We then covalently link the SOSIP to ferritin via short sortase-A linker peptides (e.g. Figure 17A). The presence of HIV-1 Env trimers on conjugated ferritin particles is confirmed with negative-stain electron microscopy.

[0133] The invention provides design of envelopes and trimer designs wherein the envelope comprises a linker which permits addition of another molecule, e.g. but not limited to a protein, such as but not limited to ferritin, or lipid, such as but not limited to cholesterol, via a Sortase A reaction. See e.g. Tsukiji, S. and Nagamune, T. (2009), Sortase-Mediated Ligation: A Gift from Gram-Positive Bacteria to Protein Engineering. ChemBioChem, 10: 787–798. doi:10.1002/cbic.200800724; Proft, T. Sortase-mediated protein ligation: an emerging biotechnology tool for protein modification and immobilisation. Biotechnol Lett (2010) 32: 1. doi:10.1007/s10529-009-0116-0; Lena Schmohl, Dirk Schwarzer, Sortase-mediated ligations for the site-specific modification of proteins, Current Opinion in Chemical Biology, Volume 22, October 2014, Pages 122-128, ISSN 1367-5931, dx.doi.org/10.1016/j.cbpa.2014.09.020; Tabata et al. Anticancer Res. 2015 Aug;35(8):4411-7; Pritz et al. *J. Org. Chem.* **2007**, *72*, 3909-3912.

[0134] The lipid modified envelopes and trimers could be formulated as liposomes. Any suitable liposome composition is contemplated.

[0135] Non-limiting embodiments of envelope designs for use in Sortase A reaction are shown in Figure 24 B-D in WO2017151801 and Figures 47 B-C in WO2017/152146, incorporated by reference in its entirety.

[0136] Additional sortase linkers could be used so long as their position allows multimerization of the envelopes.

[0137] Table 1 summary of sequences

Name	**Amino acid, nucleic acid**	**Design**	**Figure/ Note**

HV1301580_D230N_H289N_P291S;CH848.3.D1305.10.19_D949V3.DS.SOSIP_D230N_H289N_P291S (glycan hole filled)	Nt aa		12 13
>HV1301502_D1305V1;JRFL_SOSIPv6_V1_PNGS_D1305V1 (V1 loop from 10.19)	Nt aa		12 13
>HV1301405_D1305V1;CON-Schim.6R.DS.SOSIP.664_OPT_D1305V1 (V1 loop from 10.19 isolate)	Nt aa		12 13
>HV1301580_D230N_H289N_P291S;CH848.3.D1305.10.19_D949V3.DS.SOSIP_D230N_H289N_P291S (glycan holes filled)	Nt aa		12 13
>HV1301580;CH848.3.D1305.10.19_D949V3.DS.SOSIP (19CV3)	Nt aa	19CV3	12 13
>HV1301509;CH0848.3.d1305.10.19gp160	Nt aa		12 13
>HV1301503;CH848.3.D1305.10.19ch.DS.SOSIP.664	Nt aa		12 13
>HV1301504;CH848.3.D1305.10.19ch.SOSIPv6	Nt aa		12 13
>HV1301580_C_SORTA;CH848.3.D1305.10.19_D949V3.DS.SOSIP_C_SORTA	Aa nt		14 14

[0138] Table 2 Summary of mutations

Envelope	**Figure/SEQ ID No**	**V1 region**	**V3 glycosylation sites**	**UCA Ab binding**
10.17	WO2017152146 and WO2018/161049	17aa	N301 and N332	
10.17 DT	WO2017152146 and WO2018/161049	17aa N133D N138T effectively lacks	N301 and N332	DH270UCA

		glycosylation sites		
10.19	Fig. 12	17aa V1 region lacks N133 and N138 glycosylation sites	No glycosylation sites at N295, N301, N332	CH01 UCA
10.19 plus V3 loop of 10.17 (19CV3)	Fig.12, 13, Fig.14	17aa V1 region lacks N133 and N138 glycosylation sites	Add V3 regions from 10.17 has five aa difference from 10.19	CH01 UCA DH270UCA VRC26 UCA
10.19 env based with fewer than five aa changes compared to 19CV3			At least changes #2, 4, 5, and/or "GDIR" sequence	
ConS	Fig. 12, Fig. 13	17aa V1 region (from envelope 10.19) lacks N133 and N138		

[0139] Table 3. Listing non-limiting embodiments of immunogens, correlating plasmid number (see Figures 10-16) and names

Plasmid number	Protein name	Note	Figure
HV1301184	CON-S.6R.SOSIP.664		Fig. 10
HV1301185	CON-S.6R.DS.SOSIP.664		Fig. 10
HV1301186	CON-S.6R.SOSIP.664.v3.1		Fig. 10
HV1301187	CON-S.6R.SOSIP.664.v4.1		Fig. 10
HV1301188	CON-S.6R.SOSIP.664.v4.2		Fig. 10
HV1301257	CON-Schim.6R.SOSIP.664_avi		Fig. 10
HV1301258	CON-Schim.6R.DS.SOSIP.664_avi		Fig. 10
HV1301259	CON-Schim.6R.SOSIP.664v4.1_avi		Fig. 10
HV1301260	CON-Schim.6R.SOSIP.664v4.2_avi		Fig. 10
HV1301639_avi	CON-Schim.6R.DS.SOSIP.664_N130D_N135K_avi		Fig. 10
HV1301640_avi	CON-Schim.6R.DS.SOSIP.664_N138S_N141S_avi		Fig. 10
HV1301641_avi	CON-Schim.6R.DS.SOSIP.664_N130D_N135K_N138S_N141S_avi		Fig. 10
HV1301613	CON-Schim.6R.DS.SOSIP.664v4.1_OPT		Fig. 10
HV1301521_ferritin	CON-Schim.6R.DS.SOSIP.664_OPT_N130D_N135K_N138S_N141S_ferritin	I in Fig.3A	Fig. 10 Fig. 11B
	CON-Schim.6R.DS.SOSIP.664_OPT_N130X_N135X_N138X_N141X_ferritin		
HV1301521	CON-Schim.6R.DS.SOSIP.664_OPT_N130D_N135K_N138S_N141S	II in Fig.3A	Fig. 10
	CON-Schim.6R.DS.SOSIP.664_OPT_N130X_N135X_N138X_N141X		
HV1301405_N138S_N141S	CON-Schim.6R.DS.SOSIP.664_OPT_N138S_N141S	III and IV in Fig.3A	Fig. 10
	CON-Schim.6R.DS.SOSIP.664_OPT_N138X_N141X		
HV1301405	CON-Schim.6R.DS.SOSIP.664_OPT	V and VI in Fig.3A	Fig. 10
HV1301258_N301A	CON-Schim.6R.DS.SOSIP.664_N301A_avi		Fig. 10
HV1301258_N332A	CON-Schim.6R.DS.SOSIP.664_N332A_avi		Fig. 10

HV1300111_avi_N137A	CON-Sgp140CFI_avi N137A		Fig. 10
HV1300111_avi_N141A	CON-Sgp140CFI_avi N141A		Fig. 10
HV1300111_avi_V1_4Q	CON-Sgp140CFI_avi V1_4Q		Fig. 10
HV1301521_c_sorta	CON-Schim.6R.DS.SOSIP.664_OPT_N130D_N135K_N138S_N141S_C-SortaseA		Fig. 10
	CONS gp160		Fig. 16
	CONS gp160_N138X_N141X		
	CONS gp160 N130X_N135X_N138X_N141X		
HV1301521	(originally HV1301405_N130D_N135K_N138S_N141S)CON-Schim.6R.DS.SOSIP.664_OPT_N130D_N135K_N138S_N141S		Fig. 14
	>HV1301405_N138S_N141S;CON-Schim.6R.DS.SOSIP.664_OPT_N138S_N141S		Fig. 14
>HV1301405_D1305V1;	CON-Schim.6R.DS.SOSIP.664_OPT_D1305V1 (V1 loop from 10.19 isolate)		Fig. 12 Fig. 11A

[0140] Positions of mutations are HXB2 numbering and the positions which are modified in deglycosylated envelopes are bolded and underlined in ConS gp160 sequence in Figure 16.

[0141] The invention contemplates any other design, e.g. stabilized trimer, of the sequences described here in. For non-limiting embodiments of additional stabilized trimers see WO2014/042669, WO/2017151801, WO/2017152146 and WO/2018161049, all of which are incorporated by reference in their entirety, and F14 and/or Vt8 designs.

[0142] F14/Vt8 designs mutations are listed below (HXB2 numbering) with a brief explanation for each. All were originally placed in BG505 SOSIP. They were then screened via BLI of small scale transfection supernatants. From the BLI data F14, F15 and Vt8 were expressed, purified, and screened for CD4 binding and triggering.

[0143] These sets of mutations were then put into CH848 10.17 DT and CH505 M5 SOSIP (F14, Vt8, and F14+Vt8) in addition to a BG505 SOSIP F14+Vt8.

[0144] Alternative embodiments immediately follow the full sets and V3 locks below.

[0145] **Full Set** -> Pack the BMS-626529 binding site and lock the layers in place

[0146] Set of mutations referred to as **F1**: V68I, S115V, A204L, V208L, V255W, N377L, M426W, M434W, H66S.

[0147] Eliminate* N377L, M426W, and M434W -> Avoid over-packing the area - N377 may be important for folding (is not totally buried). Eliminate means, that F2 construct includes all F1 mutations except N337L, M426W, and M434W.

[0148] Set of mutations referred to as **F2**: V68I, S115V, A204L, V208L, V255W, H66S

[0149] Eliminate S115V -> Adding a V may be too large for the area

[0150] Set of mutations referred to as **F3**: V68I, A204V, V208L, V255L, H66S

[0151] Eliminate A204V -> Adding a V may be too large for the packed region A204 resides (adding E causes opening of the apex)

[0152] Set of mutations referred to as **F4**: V68I, S115V, V208L, V255L, H66S

[0153] Retain N377L for minimal set -> Above tested N377L elimination from full set, test here whether N377L stabilizes

[0154] Set of mutations referred to as **F5**: V68I, S115V, A204L, V208L, V255W, N377L, H66S

[0155] Add W69L to minimal set -> previous work suggests aromatic residues in position 69 are destabilizing - test here

[0156] Set of mutations referred to as **F6**: V68I, S115V, A204L, V208L, V255L, W69L

[0157] Use W69V instead of W69L -> test whether side chain length alters potential stabilizing effect

[0158] Set of mutations referred to as **F7**: V68I, S115V, A204L, V208L, V255L, W69V

[0159] Use W69A instead of W69L/V -> further test whether side chain length alters potential stabilizing effect

[0160] Set of mutations referred to as **F8**: V68I, S115V, A204L, V255L, V208L, W69A

[0161] Reintroduce M426W -> test a minimally reduced set - effect of M's

[0162] Set of mutations referred to as **F9**: V68I, S115V, A204L, V208L, V255W, N377L, M426W, H66S

[0163] Reintroduce M434W -> test a minimally reduced set - effect of M's

[0164] Set of mutations referred to as **F10**: V68I, S115V, A204L, V208L, V255W, N377L, M434W, H66S

[0165] Introduce additional H72 mutation -> can P favor loop turn stabilizing TRP69 Loop in W bound state

[0166] Set of mutations referred to as **F11**: V68I, S115V, A204V, V208L, V255L, H72P, H66S

[0167] Test minimal set with H66K rather than S -> is charge a better solution to polar switch

[0168] Set of mutations referred to as **F12**: V68I, S115V, V208L, V255L, H66K

[0169] Eliminate H66S from F1 -> H66 may be important for loop configuration

[0170] Set of mutations referred to as **F13**: V68I, S115V, A204L, V208L, V255W, N377L, M426W, M434W

[0171] Minimal Set 2 -> Eliminate H66 and swap S115V for A204V; H66 could be important for loop and A204 my better stabilize that S115V

[0172] Set of mutations referred to as **F14**: V68I, A204V, V208L, V255L

[0173] Minimal Set 3 -> Add N377L to test for further stabilization

[0174] Set of mutations referred to as **F15**: V68I, A204L, V208L, V255W, N377L

[0175] **V3 lock - Full Set**

[0176] Set of mutations referred to as **Vt1**: Y177F, T320L, D180A, Q422L, Y435F, Q203M, E381L, R298M, N302L, N300L

[0177] Eliminate R298M and E381L -> Determine whether these two are stabilizing rather than destabilizing

[0178] Set of mutations referred to as **Vt2**: Y177F, T320L, D180A, Q422L, Y435, Q203M, N302L, N300L

[0179] Eliminate E381L -> Determine whether this residue is required to stabilize R298

[0180] Set of mutations referred to as **Vt3**: Y177F, T320L, D180A, Q422L, Y435, Q203M, R298M, N302L, N300L

[0181] Eliminate R298M -> Determine whether this reside stabilizes E381

[0182] Set of mutations referred to as **Vt4**: Y177F, T320L, D180A, Q422L, Y435, Q203M, E381L, N302L, N300L

[0183] Retain Y177 and Y435 -> May stabilize interior through H-bonding

[0184] Set of mutations referred to as **Vt5**: T320L, D180A, Q422L, Q203M, E381L, R298M, N302L, N300L

[0185] Retain Y177 and Y435 while eliminating R298 and E381 mutations -> A minimal set avoiding possible problems from charged pair mutations

[0186] Set of mutations referred to as **Vt6**: T320L, D180A, Q422L, Q203M, N302L, N300L

[0187] Dennis Burton Set -> Control for comparison

[0188] Set of mutations referred to as **Vt7**: R298A, N302F, R304V, A319Y, T320M

[0189] Eliminate D180A -> D180 appears to be destabilizing but may be stabilizing

[0190] Set of mutations referred to as **Vt8**: T320M, Q422M, Q203M, N302L, N300L

[0191] Add S174V -> S174 is on the periphery but may be stabilizing with a hydrophobe

[0192] Set of mutations referred to as **Vt9**: T320M, Q422M, Q203M, N302L, N300L, S174V

[0193] Set (DS-SOSIP.4mut) -> Additional Control Set

[0194] Set of mutations referred to as **Vt10**: I201C, A443C, L154M, N300M, N302M, T320L

[0195] *In the above description, "eliminate" means that full set number "N" construct includes all full set number "N-1" mutations except the mutations identified as eliminated.

[0196] Contemplated also are subsets of the mutations within a set. In a non-limiting embodiment, the mutations in Set F14 could be further parsed out to determine if there are fewer mutations or combinations of fewer mutations than in Set 14 which provide stabilization of the trimer.

[0197] In certain embodiments the invention provides an envelope comprising 17aa V1 region without N133 and N138 glycosylation, and N301 and N332 glycosylation sites, and further comprising "GDIR" motif see Ex. 1, wherein the envelope binds to UCAs of V1V2 Abs and V3 Abs.

EXAMPLES

[0198] The following specific examples are to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever. Without further elaboration, it is believed that one skilled in the art can, based on the description herein, utilize the present invention to its fullest extent.

Example 1: Pan-bnAb-engaging Immunogens

[0199] **Example 1A:** This example describes design of HIV-1 envelopes antigenic for cross-epitope bnAb UCAs.

[0200] The discovery of broadly neutralizing antibodies (bnAbs) in HIV-1 infected individuals has provided evidence that the human immune system can target highly conserved epitopes on HIV-1 envelope. However, bnAbs have not been reproducibly induced with a vaccine, in primates. One approach to improve the induction of bnAbs is to specifically design immunogens that bind to the precursor B cell that gives rise to the bnAb. While highly affinity matured HIV-1 bnAbs react with many Envelope proteins, their precursors bind only to select Envs. Currently, immunogens exist that can bind to a single bnAb precursor. These Envs have the disadvantage of relying on a single bnAb precursor to be present in most individuals. If the bnAb precursor antibody is not present in that individual then the vaccine will not have the intended effect of inducing a specific type of antibody response. To improve the chances that an individual has the bnAb precursor that can engage the vaccine immunogen, we created a vaccine immunogen that can bind to multiple bnAb

precursors. We designed the immunogen to interact with bnAbs precursors that interact with the first and second variable loop and glycans proximal to this loop—an epitope called V1V2-glycan. Secondly, the immunogen was also designed to interact with a bnAb precursor that bound to the third variable region and surrounding glycans on HIV-1 envelope—the V3-glycan site.

[0201] The immunogen was designed by creating a chimera of two HIV-1 envelope sequences that were derived from the HIV-1 infected individual CH0848 (See WO/2017152146 and WO/2018161049 references). The first Env CH0848.3.D0949.10.17 is antigenic for V3-glycan antibodies and was selected because it had a short first variable region in Env and bound to a V3-glycan antibody that possessed only 5 mutations (Bonsignori et al STM 2017). We modified this Env by removing glycosylation sites at 133 and 138 and found V3-glycan antibodies bound better to the Env when the glycosylation site was removed. These two glycosylation sites were identified as inhibitory in a neutralization screen where glycosylation sites on Env were removed to determine which glycans were required for neutralization by V3-glycan antibodies. For the CH0848.3.D0949.10.17 envelope we removed the glycosylation by substituting asparagine for amino acids that normally occur at positions 133 and 138 in other viruses. This glycan-modified Env bound with low nanomolar affinity to the V3-glycan bnAb precursor DH270 UCA3. To determine if a similar Env may have been present in the infected individual and could have potentially initiated the V3-glycan lineage in vivo, we screened all of the autologous virus sequences isolated from the infected individual CH0848 for viruses with a 17 amino acid variable region 1 and no glycans within the variable region except at position 156. We identified two sequences, with these characteristics. The first sequence CH0848.3.D1305.10.19 was produced as a recombinant protein. In biolayer interferometry assays it did not bind to V3-glycan antibodies. We created a pseudovirus expressing this Env and also found that V3 glycan antibodies did not neutralize it. However, we found that V1V2-glycan antibodies could bind to the recombinant protein. This was in contrast to CH0848.3.D0949.10.17, which lacked binding to V1V2-glycan bnAbs and precursors but was antigenic for V3-glycan antibodies. We inspected the sequences of the V1V2 and V3 regions and found that CH0848.3.D1305.10.19 lacked three glycans at positions 295, 301, and 332 usually bound by V3-glycan antibodies. To restore these V3 proximal glycosylation sites in CH0848.3.D1305.10.19 we used the V3 sequence of CH0848.3.D0949.10.17—the new envelope referenced as 19CV3. The modification of the CH0848.3.D1305.10.19 sequence to 19CV3 resulted in the addition of glycosylation sites at positions 301 and 332. We again

made a recombinant protein of the chimeric envelope and found it bound to V1V2-glycan bnAbs as well as V3-glycan bnAbs—a combination of the phenotypes of the two parental envelopes. We next tested the binding of the bnAb precursors for V1V2 and V3-glycan sites. We found that 19CV3 bout to the bnAb precursor for two V1V2 glycan bnAb, CH01 and VRC26, and V3 glycan Ab DH270.

[0202] With reference to CH0848 10.17DT SOSIP sequence, see WO2018/161049, incorporated by reference in its entirety.

[0203] For non-limiting examples of hole-filled CH848 703010848.3.d0949.10.17 envelopes see WO/2017152146 and WO2018/161049, inter alia without limitation, Figures 44A-D, incorporated by reference in their entirety.

[0204] The immunogens of the invention can be delivered by any suitable mechanism.

[0205] In non-limiting embodiments, these could be Adeno-associated virus (AAV) vectors; Non-replicating viral vectors; vectors which provide sustained expression of the immunogen;

[0206] Vectors which can transduce dendritic cells, which present transgene(immunogen) in complex with MHCII to naïve T cells. Constant antigen production could lead to improved clonal persistence, enhanced germinal center reactions, and higher somatic mutation;

[0207] In certain embodiments, the immunogens could be multimerized.

[0208] **Example 1B**--This example describes design of HIV-1 envelopes antigenic for cross-epitope bnAb UCAs--ConS envelope designs as panbnAb immunogens

[0209] To cover the diversity of HIV-1 isolates that circulate globally a consensus envelope was derived from all group M HIV-1 isolates available at the time (Liao HX et al Virology. 2006 Sep 30; 353(2): 268–282., See also US Patent No. 8,071,107 and all parent applications and application claiming priority to) called CON-S. To induce neutralizing antibodies it is hypothesized that the immunogen should mimic the native, fusion-competent envelope on viruses. To create stable mimics of the HIV-1 Env CON-S we created SOSIP gp140s. The SOSIP gp140 was stabilized by introducing BG505 amino acids into the gp120 and gp41 regions as we have described previously (Saunders KO, Vercokzy L et al. Cell Reports. Volume 21, ISSUE 13, P3681-3690, December 26, 2017). The Env was further stabilized by introducing a disulfide bond between amino acids at position 201 and 433 (Do-Kwon Y et al Nat Struct Mol Biol. 2015 Jul;22(7):522-31. doi: 10.1038/nsmb.3051. Epub 2015 Jun 22.).

[0210] The CON-S sequence was furthered optimized to bind to antibodies that target the V3-glycan broadly neutralizing site by removing glycans that were determined in neutralization assays to inhibit V3-glycan antibody binding and neutralization. We hypothesize that broadly neutralizing antibody precursors have low affinity for HIV-1 Env

which necessitates reducing steric barriers and glycosylation changes that hinder precursor antibody binding. In neutralization assays we identified that glycans attached between N131 and N141 prohibited neutralization by precursor antibodies that were developing neutralization breadth.

[0211] To improve binding to the V3-glycan site on CON-S stabilized gp140 SOSIPs we removed glycosylation sites at 130, 135, 138, and 141 by substituting asparagine for naturally occurring amino acids identified in the HIV-1 sequence database. The mutant Env contained N130D, N135K, N138S, and N141S mutations. Using mass spectrometry we verified that the glycans at 295, 301, and 332 were still the high mannose glycans preferentially bound by broadly neutralizing antibodies PGT128, PGT124, PGT135, DH270, BF520, and BG18. While removal of the V1 glycans may allow better binding to Env, the affinity for Env may be low for certain V3-glycan bnAb precursors. It has been shown that B cell receptors recognize low affinity antigen better when it is presented on a surface rather than free in solution (Batista F and Neuberger M J EMBO 2000, 19(4):513-520). Thus, we took the Env and arrayed it on the surface of a ferritin nanoparticle so that 8 copies of the CON-S SOSIP trimer could be displayed to B cells to maximize avidity of the BCR:SOSIP interaction. In total, a stabilized soluble HIV-1 Env trimer was derived from a consensus of group M and inhibitory glycans were removed to promote V3-glycan bnAb precursor binding. The optimized Env was arrayed on ferritin nanoparticles to enhance avidity between Env and B cell receptors.

[0212] The removal of four glycans in the V1 loop was hypothesized to permit binding of Env to unmutated bnAb precursors to initiate bnAb lineages. To select the bnAb intermediate antibodies within a lineage that are acquiring the ability to bind to multiple native Envs, we created a CON-S SOSIP Env trimers that added back the N130 and N135 glycosylation sites. This Env lacks glycosylation sites at 138 and 141 functions to select the antibodies that bind to Env with the correct mode to accommodate the N130 and N135 glycans. In a sequential vaccine this Env would be administered after 4 glycan deleted Env but before the wildtype Env so that glycans are sequentially added back to the Env to select the small population of B cells that recognize the V3-glycan site with the correct binding orientation.

Example 2

[0213] Glycan-optimized trimeric HIV-1 envelope elicits glycan-dependent autologous tier 2 neutralizing antibodies in rhesus macaques (See Figures 1-6, 18 et seq)

[0214] This example is based on the hypothesis that: Nanoparticle immunogens are necessary to overcome the low affinity between V3-glycan bnAb precursors and HIV-1 Env; HIV-1 Env should be enriched for Man9GlcNAc2 in order to optimally engage V3-glycan bnAb precursors; V1 glycans are inhibitory for early intermediate antibodies, thus sequential selection of antibodies that can accommodate V1 glycans will be necessary.

[0215] Introduction: Vaccine elicitation of broadly neutralizing antibodies (bnAbs) against HIV-1 has yet to be achieved. The target of bnAbs is HIV-1 envelope (Env) which is shielded by host glycans that hinder its recognition by antibodies. During natural infection, bnAbs develop that recognize the glycans and peptide proximal to the third variable region (V3-glycan). These glycan-dependent antibodies are protective in nonhuman primate models of HIV-1 infection. We previously observed that reactivity with Env was enhanced for V3-glycan bnAbs when the Env was enriched for Man9GlcNAc2 glycans or when V1 glycans were removed. We hypothesize that glycan-dependent bnAbs can be induced in primates with a vaccine if the immunogens are optimized to engage V3-glycan bnAb precursors and subsequently select for B cells within those lineages that are developing neutralization breadth.

[0216] The scientific premises of the non-human primate study (NHP145) is V3 glycan precursors prefer kif treated Env; A multimer is needed to activate the germline precursors because the affinity is so low; V3 glycan precursors have to learn to accommodate processed glycans one at a time

[0217] Methods: Recombinant trimeric HIV-1 CON-S Env was made as a SOSIP trimer and arrayed on ferritin nanoparticles. To enrich for Man9GlcNAc2 some Env were treated with kifunensine (kif). Trimer formation was determined by negative stain electron microscopy (EM). Antigenicity of the Envs was determined by Bio-layer interferometry. Four rhesus macaques were vaccinated 6 times with a series of HIV-1 Env glycosylation variants optimized to be antigenic for V3-glycan bnAbs as shown in Fig. 2. Binding and neutralizing antibodies were measured by ELISA and the TZM-bl assay respectively.

[0218] Animal study: Four non-human primates (NHPs) were immunized with the immunization regiment shown in Figure 3A, Figure 6. Recombinant protein was administered at a dose of 100microgr in TLR4 adjuvant at every 6 weeks.

[0219] Conclusions:

[0220] Modified CON-S nanoparticles bind to the precursors of V3-glycan and V1V2 glycan bnAbs.

[0221] Multimerization of HIV-1 Env induces more durable antibody responses than free trimer.

[0222] Neutralizing antibody responses show that vaccination can elicit glycan-dependent neutralizing antibodies against the same Asn301 glycan targeted by bnAbs.

[0223] This example showed that CON-S SOSIP nanoparticle is antigenic for V3-glycan bnAb precursors. It also showed selection of sequential CON-S SOSIPs with glycan modifications boost glycan antibodies. The examples showed that this immunization regimen elicited autologous tier 2 neutralizing antibodies that did not target a glycan hole near N362. Autologous tier 2 neutralizing antibodies were N301 glycan–dependent in 2 of 4 macaques. The N301 glycan-dependent antibodies were distinct from DH501 in that they neutralized untreated CON-S, and thus did not require Man9GlcNAc2 enrichment.

[0224] References:

[0225] Stewart-Jones et al. Cell. 2016 May 5;165(4):813-26. doi: 10.1016/j.cell.2016.04.010. Epub 2016 Apr 21.

[0226] Saunders et al., 2017 Cell Rep., 18 (2017), pp. 2175-2188.

[0227] Antibodies will be isolated (e.g. by single cell sorting), cloned and further analyzed for their properties including binding to autologous and heterologous envelopes, neutralization, etc. The goal is to determine types and specificities of induced antibodies, and whether any broad neutralizing or otherwise protective antibodies lineages are introduced.

[0228] Additional NHP studies could be conducted to determine whether immunization provides protective responses.

Example 3

[0229] Con-S V1 delta glycans were also teste in Ca2+ flux assay.

[0230] Figures 18A-18C show CON-S envelope induction of B cell receptor signaling in Ramos B cell lines expressing HIV-1 broadly neutralizing antibodies. The 3 antibodies are from three different points of maturation of the DH270 bnAb B cell lineage. In 18A the CON-S envelope inducing B cell receptor signaling in cells expressing the first intermediate antibody (DH270 IA4) from the DH270 lineage as well as a broadly neutralizing antibody (DH270) from the same lineage. They demonstrate that the envelope is antigenic for the earliest intermediate antibody within the DH270 lineage. In 18C, the presence of glycans in V1 of CON-S abrogates binding to DH270 IA4. The effect of glycan removal is not the same

for another envelope JR-FL. The removal of glycans in V1 of JRFL is not sufficient to confer binding to the DH270 IA4 antibody.

[0231] Various recombinant proteins, trimers and/or nanoparticles were purified by chromatography, including antibody affinity chromatography (e.g. PGT145).

[0232] 442EML(b)
CON_Schim.6R.DS.SOSIP.664_OPT_N130D_N135K_N138S_N141S_ferritin 5uM-Kif 293F

[0233] 455EML
CONSchim.6R.DS.SOSIP.664_OPT_N130D_N135K_N138S_N141S_ferritin 293F

[0234] Antigenicity of recombinant ConS envelopes is shown in Figures 30-34.

What is claimed is

1. A recombinant CON-S HIV-1 envelope comprising a V1 region, wherein the envelope lacks glycosylation at positions N138 and N141 (HXB2 numbering and Figure 16).
2. A recombinant CON-S HIV-1 envelope comprising a V1 region, wherein the envelope lacks glycosylation at positions N130, N135, N138, and N141 (HXB2 numbering and Figure 16).
3. A recombinant CON-S HIV-1 envelope comprising a 17aa-long V1 region, wherein the envelope lacks glycosylation at positions N138 and N141 (HXB2 numbering).
4. A recombinant CON-S HIV-1 envelope comprising a 17aa-long V1 region, wherein the envelope lacks glycosylation at positions N130, N135, N138, and N141 (HXB2 numbering).
5. The recombinant envelope of claim 1, wherein the envelope comprises all consecutive amino acids after the signal peptide of CON-Schim.6R.DS.SOSIP.664_OPT_N138X_N141X (Table 3).
6. The recombinant envelope of claim 1, wherein the envelope comprises all consecutive amino acids after the signal peptide of CON-Schim.6R.DS.SOSIP.664_OPT_N138S_N141S (Table 3, Figure 10).
7. The recombinant envelope of claim 2, wherein the envelope comprises all consecutive amino acids after the signal peptide of CON-Schim.6R.DS.SOSIP.664_OPT_N130X_N135X_N138X_N141X (Table 3).
8. The recombinant envelope of claim 2, wherein the envelope comprises all consecutive amino acids after the signal peptide of CON-Schim.6R.DS.SOSIP.664_OPT_N130D_N135K_N138S_N141S (Table 3, Figure 10).
9. The recombinant envelope of claim 2, wherein the envelope comprises all consecutive amino acids after the signal peptide of CON-Schim.6R.DS.SOSIP.664_OPT_N130X_N135X_N138X_N141X_ferritin (Table 3).
10. The recombinant envelope of claim 2, wherein the envelope comprises all consecutive amino acids after the signal peptide of CON-Schim.6R.DS.SOSIP.664_OPT_N130D_N135K_N138S_N141S_ferritin (Table 3, Figure 10, Figure 11B).

11. The recombinant envelope of claim 3, wherein the envelope comprises all consecutive amino acids after the signal peptide of CON-Schim.6R.DS.SOSIP.664_OPT_D1305V1 (Table 1, Figure 13).

12. The recombinant CON-S envelope of any one of the preceding claims wherein the envelope is a protomer comprised in a stable trimer.

13. The recombinant envelope of any of claims 1-12, wherein the envelope is Man9GlcNAc-enriched.

14. A composition comprising the recombinant CON-S envelopes of any one of claims 1-13.

15. The composition of claim 14, wherein the composition comprises a carrier.

16. A nucleic acid encoding the recombinant envelopes of any one of claim 1-12.

17. A composition comprising the nucleic acid of claim 16 and a carrier.

18. The composition of claim 17, wherein the nucleic acid is a modified mRNA.

19. A composition comprising the recombinant CON-S envelope of any one of claim 1-13, wherein the recombinant CON-S envelope is multimerized and wherein optionally in some embodiments the envelope is comprised in a nanoparticle.

20. The composition of claim 19, wherein the recombinant CON-S envelope is comprised in a nanoparticle which is a ferritin nanoparticle.

21. A method of inducing an immune response in a subject comprising administering an immunogenic composition comprising the recombinant CON-S envelopes of any one of claim 1-13 or the composition of any one of claims 14, 15, or 18.

22. The method of claim 21, wherein the immunogenic composition comprises a first immunogen comprising all consecutive amino acids after the signal peptide of CON-Schim.6R.DS.SOSIP.664_OPT_N130D_N135K_N138S_N141S_ferritin (Table 3, Figure 10, Figure 11B), and wherein the immunogen is optionally Man9GlcNAc-enriched, optionally multimerized in a nanoparticle, wherein the nanoparticles is a ferritin nanoparticle.

23. The method of claim 22, further comprising administering a second immunogenic composition comprising a second immunogen comprising all consecutive amino acids after the signal peptide of CON-Schim.6R.DS.SOSIP.664_OPT_N130D_N135K_N138S_N141S (Table 3, Figure 10, Figure 11B), and wherein the immunogen is optionally multimerized in a trimer.

24. The method of claim 23, further comprising administering a third immunogenic composition comprising a third immunogen comprising all consecutive amino acids after the signal peptide of CON-Schim.6R.DS.SOSIP.664_OPT_N138S_N141S (Table 3, Figure 10), wherein the immunogen is optionally Man9GlcNAc-enriched and optionally multimerized in a trimer.
25. The method of claim 24, further comprising administering a fourth immunogenic composition comprising a fourth immunogen comprising all consecutive amino acids after the signal peptide of CON-Schim.6R.DS.SOSIP.664_OPT_N138S_N141S (Table 3, Figure 10), wherein the immunogen is optionally multimerized in a trimer.
26. The method of claim 25, further comprising administering a fifth immunogenic composition comprising a fifth immunogen comprising all consecutive amino acids after the signal peptide of CON-Schim.6R.DS.SOSIP.664_OPT_ (Table 3, Figure 10), wherein the immunogen is optionally Man9GlcNAc-enriched and optionally multimerized in a trimer.
27. The method of claim 26, further comprising administering a sixth immunogenic composition comprising a sixth immunogen comprising all consecutive amino acids after the signal peptide of CON-Schim.6R.DS.SOSIP.664_OPT_ (Table 3, Figure 10), wherein the immunogen is optionally multimerized in a trimer.
28. The method any one of claims 21-27, further comprising administering an adjuvant.
29. The method any one of claims 21-27, wherein the composition is administered as a prime and/or a boost, and optionally wherein the composition comprises nanoparticles.

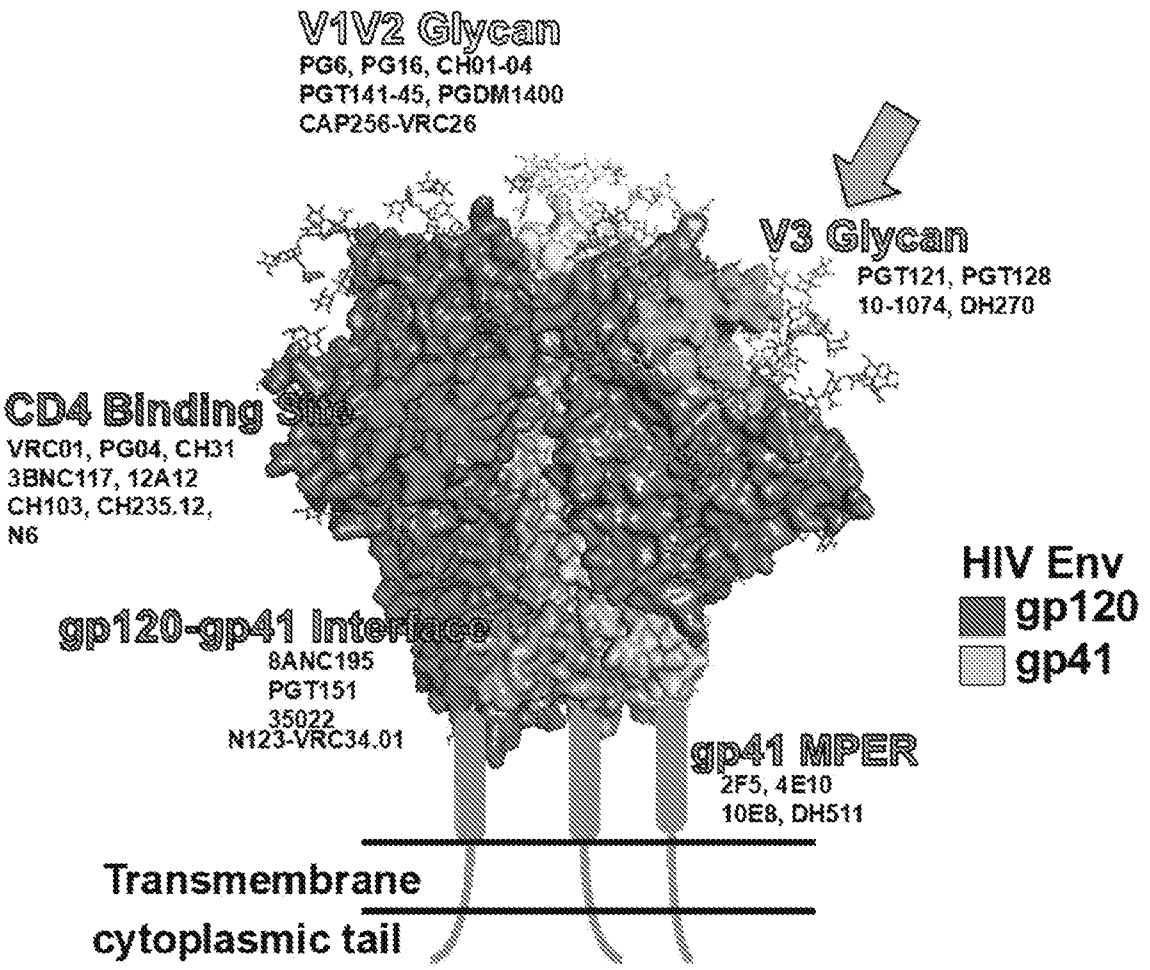

V1V2 Glycan
PG6, PG16, CH01-04
PGT141-45, PGDM1400
CAP256-VRC26
V3 Glycan
PGT121, PGT128
10-1074, DH270
CD4 Binding Site
VRC01, PG04, CH31
3BNC117, 12A12
CH103, CH235.12,
N6
gp120-gp41 Interface
8ANC195
PGT151
35022
N123-VRC34.01
gp41 MPER
2F5, 4E10
10E8, DH511
HIV Env
gp120
gp41
Transmembrane
cytoplasmic tail

Figure 1A

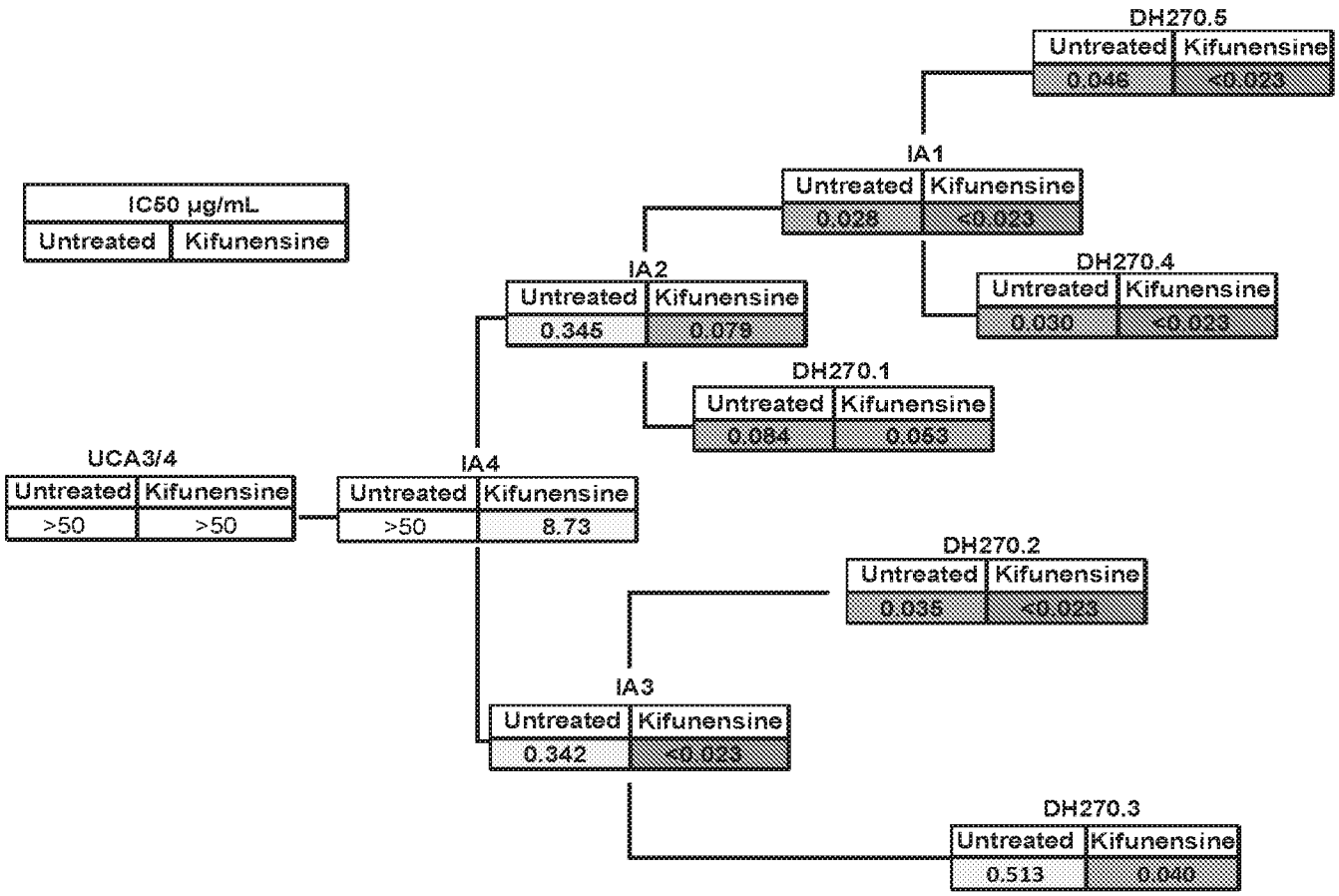

IC50 µg/mL
Untreated
Kifunensine
UCA3/4
Untreated
Kifunensine
>50
>50
IA4
Untreated
Kifunensine
>50
8.73
IA2
Untreated
Kifunensine
0.345
0.079
IA1
Untreated
Kifunensine
0.028
<0.023
DH270.5
Untreated
Kifunensine
0.046
<0.023
DH270.4
Untreated
Kifunensine
0.030
<0.023
DH270.1
Untreated
Kifunensine
0.084
0.053
DH270.2
Untreated
Kifunensine
0.035
<0.023
IA3
Untreated
Kifunensine
0.342
<0.023
DH270.3
Untreated
Kifunensine
0.513
0.040

Figure 1B

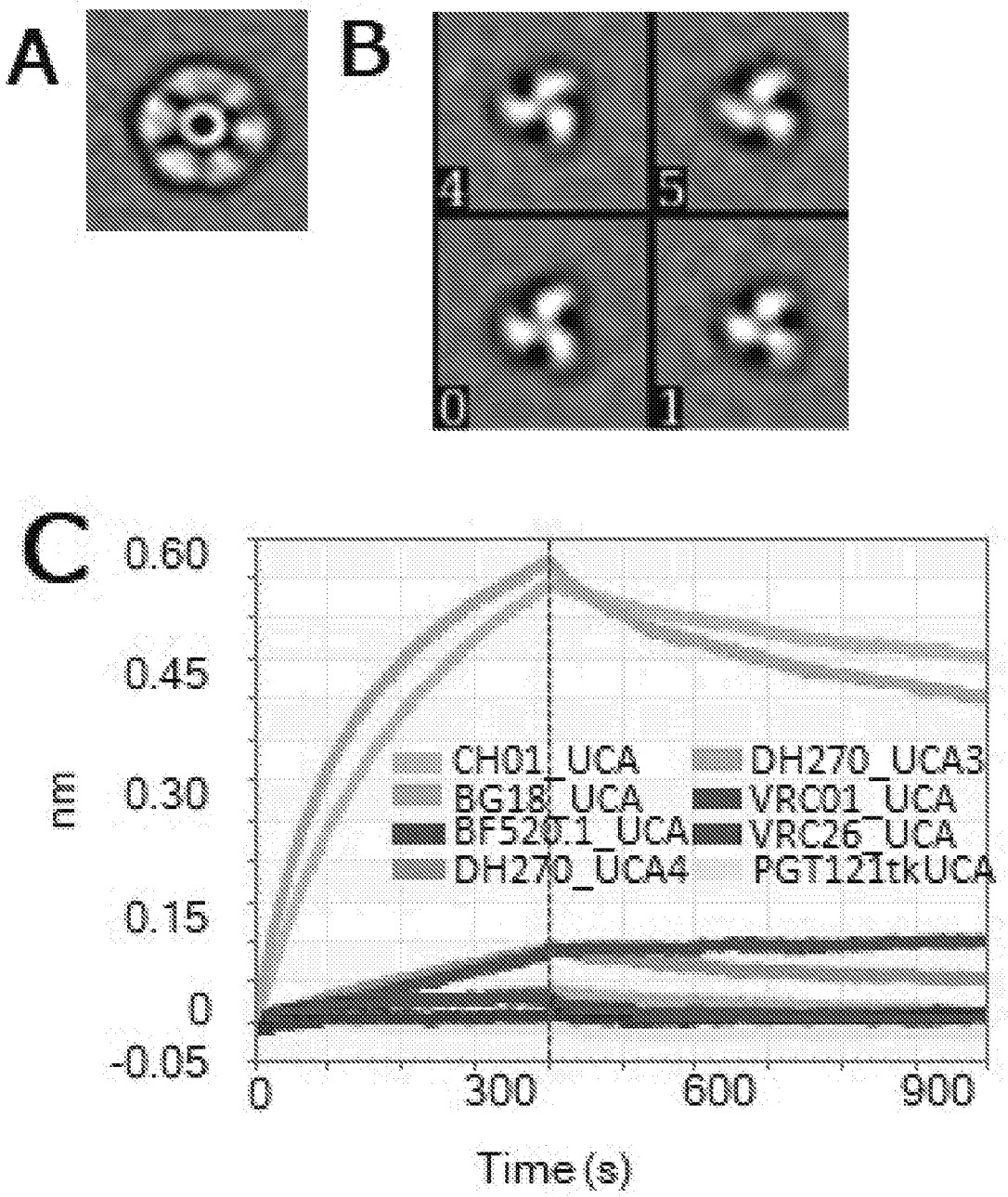
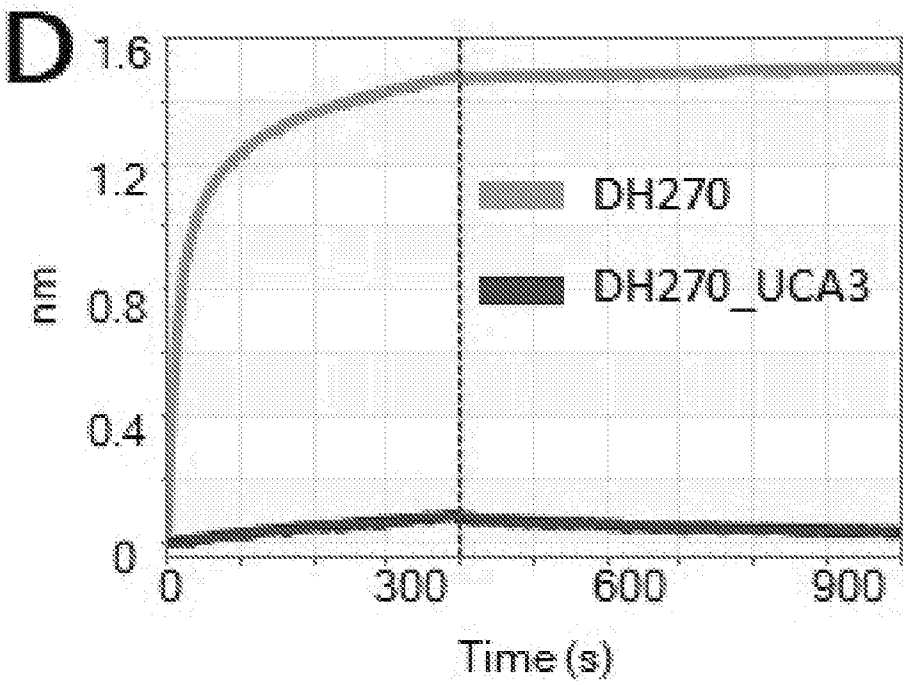


Figure 2A, 2B and 2C



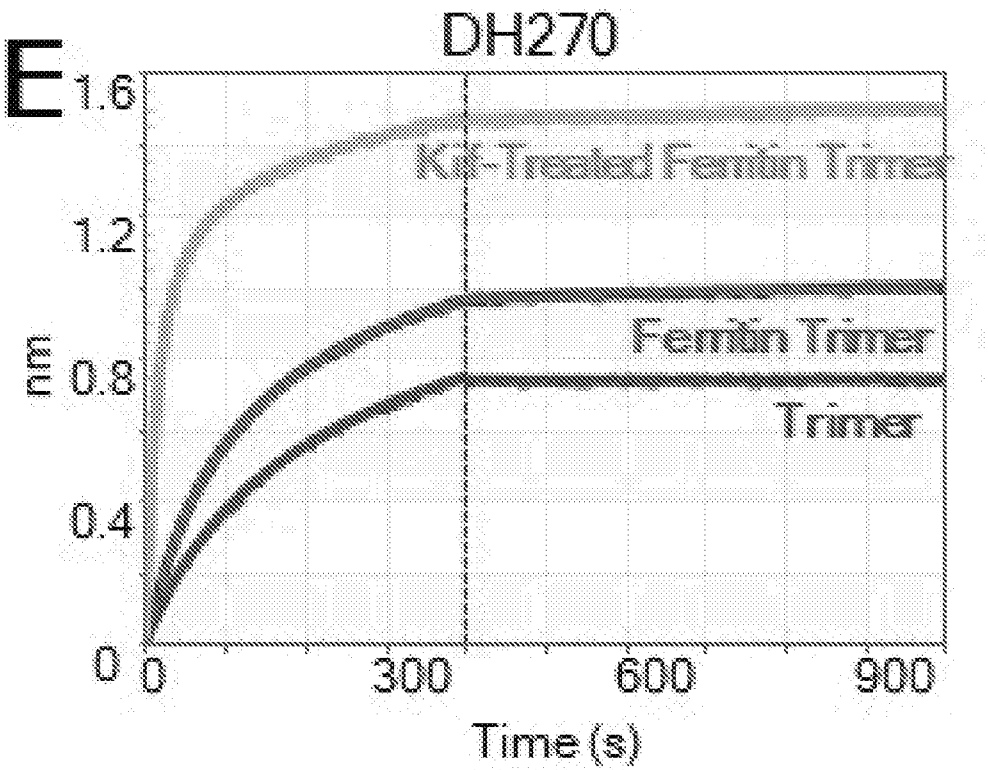


Figure 2D and 2E

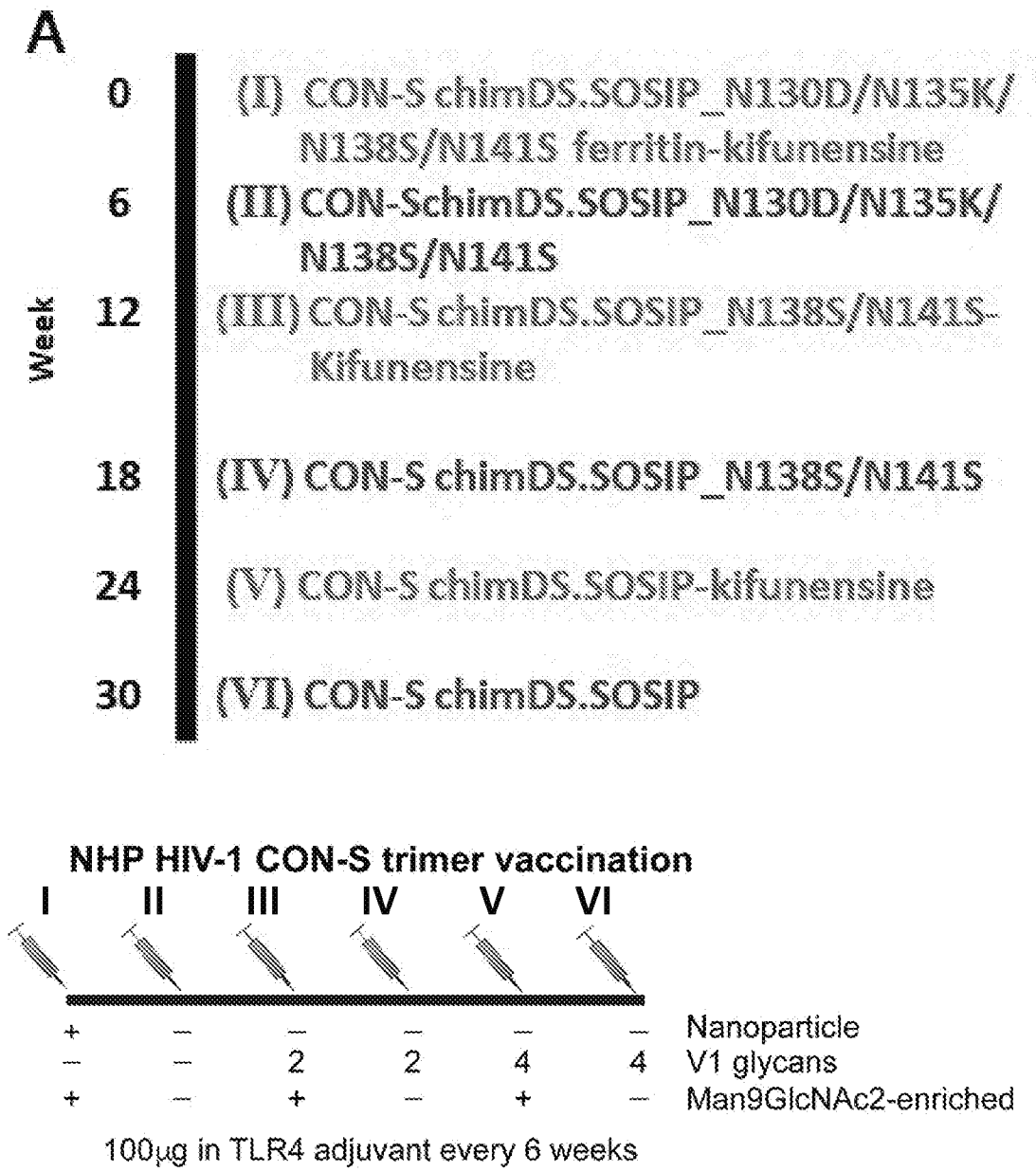

A
Week
0
6
12
18
24
30
(I) CON-S chimDS.SOSIP_N130D/N135K/ N138S/N141S ferritin-kifunensine
(II) CON-SchimDS.SOSIP_N130D/N135K/ N138S/N141S
(III) CON-S chimDS.SOSIP_N138S/N141S-Kifunensine
(IV) CON-S chimDS.SOSIP_N138S/N141S
(V) CON-S chimDS.SOSIP-kifunensine
(VI) CON-S chimDS.SOSIP
NHP HIV-1 CON-S trimer vaccination
I II III IV V VI
+ – – – – – Nanoparticle
– – 2 2 4 4 V1 glycans
+ – + – + – Man9GlcNAc2-enriched
100μg in TLR4 adjuvant every 6 weeks

Figure 3A

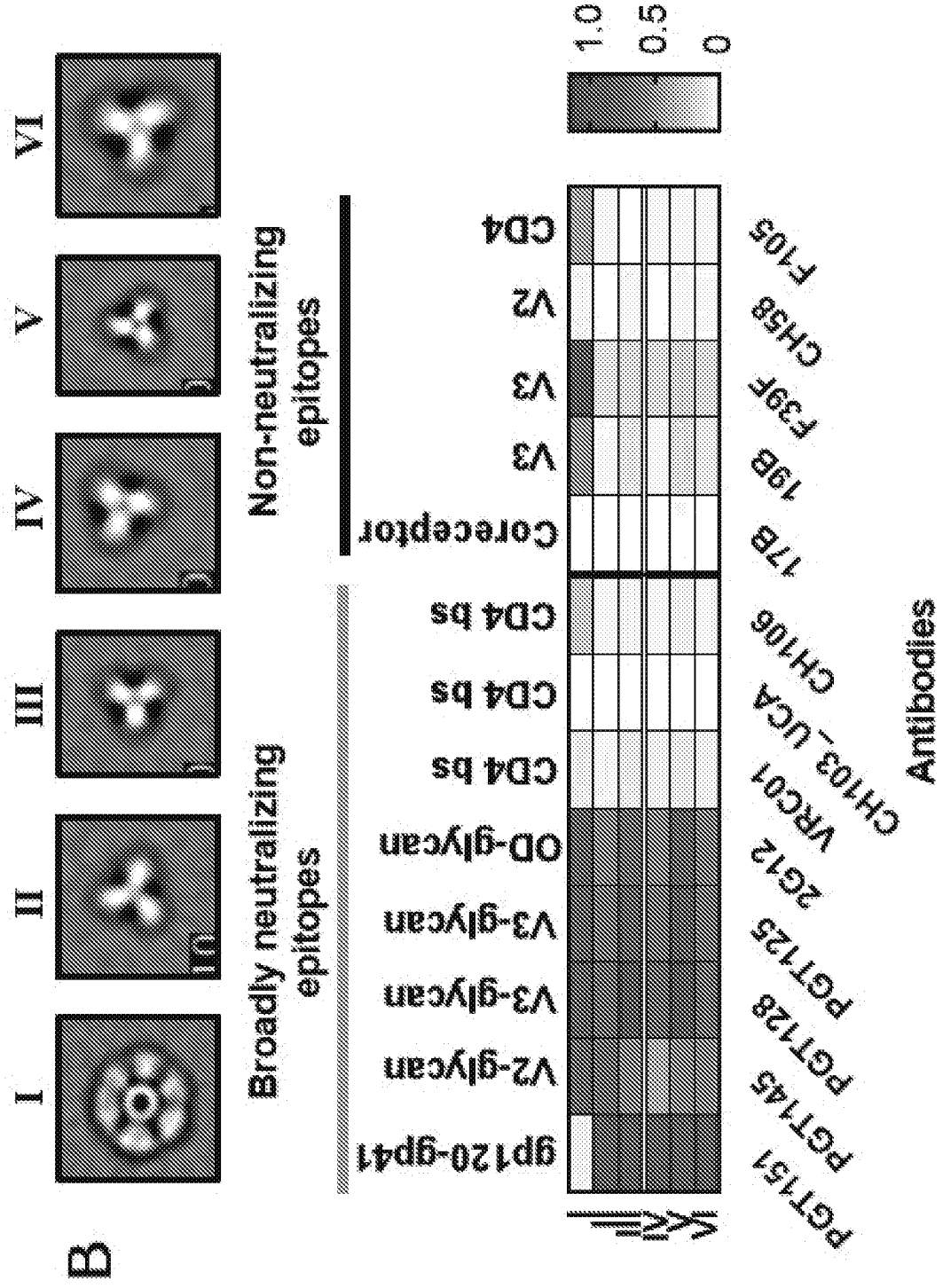

B
I
II
III
IV
V
VI
Broadly neutralizing epitopes
Non-neutralizing epitopes
gp120-gp41
V2-glycan
V3-glycan
V3-glycan
OD-glycan
CD4 bs
CD4 bs
CD4 bs
Coreceptor
V3
V3
V2
CD4
PGT151
PGT145
PGT128
PGT125
2G12
VRC01
CH103_UCA
CH106
17B
19B
F39F
CH58
F105
1.0
0.5
0
Antibodies

Figure 3B

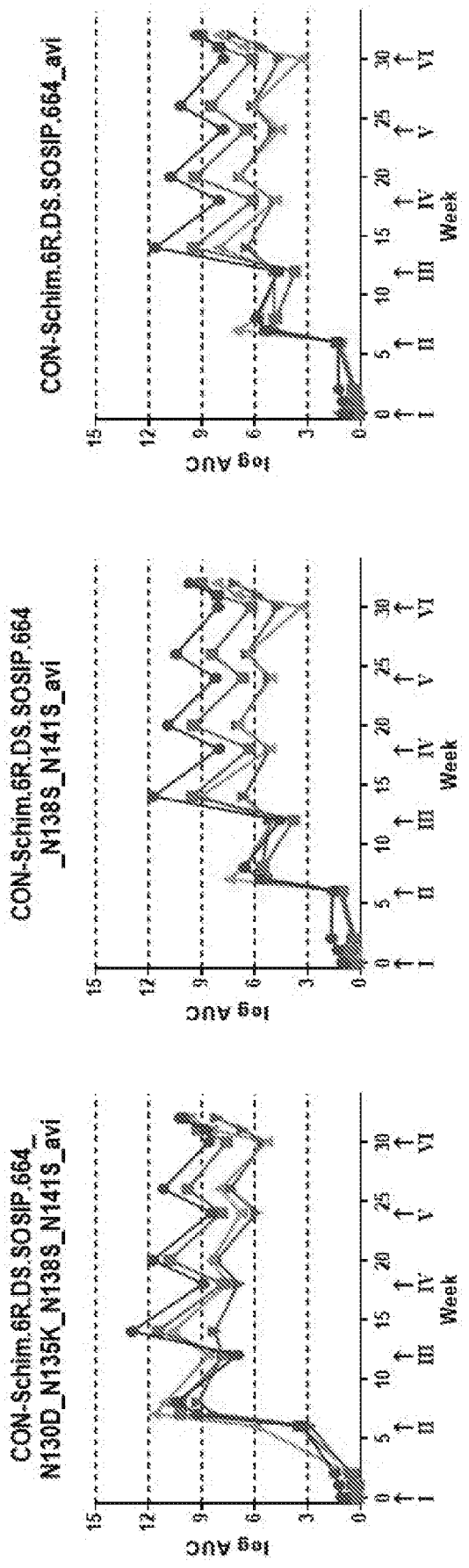

CON-Schim.6R.DS.SOSIP.664_avi
CON-Schim.6R.DS.SOSIP.664
_N138S_N141S_avi
CON-Schim.6R.DS.SOSIP.664_
N130D_N135K_N138S_N141S_avi
log AUC
Week
I
II
III
IV
V
VI

Figure 3C

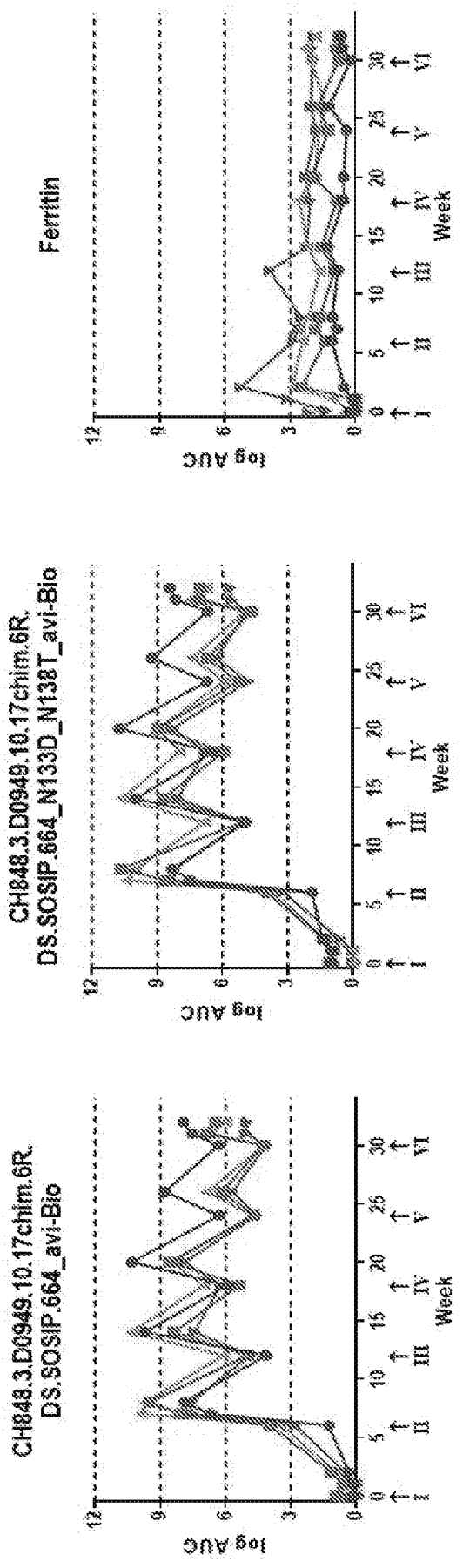


Figure 3C (cont)

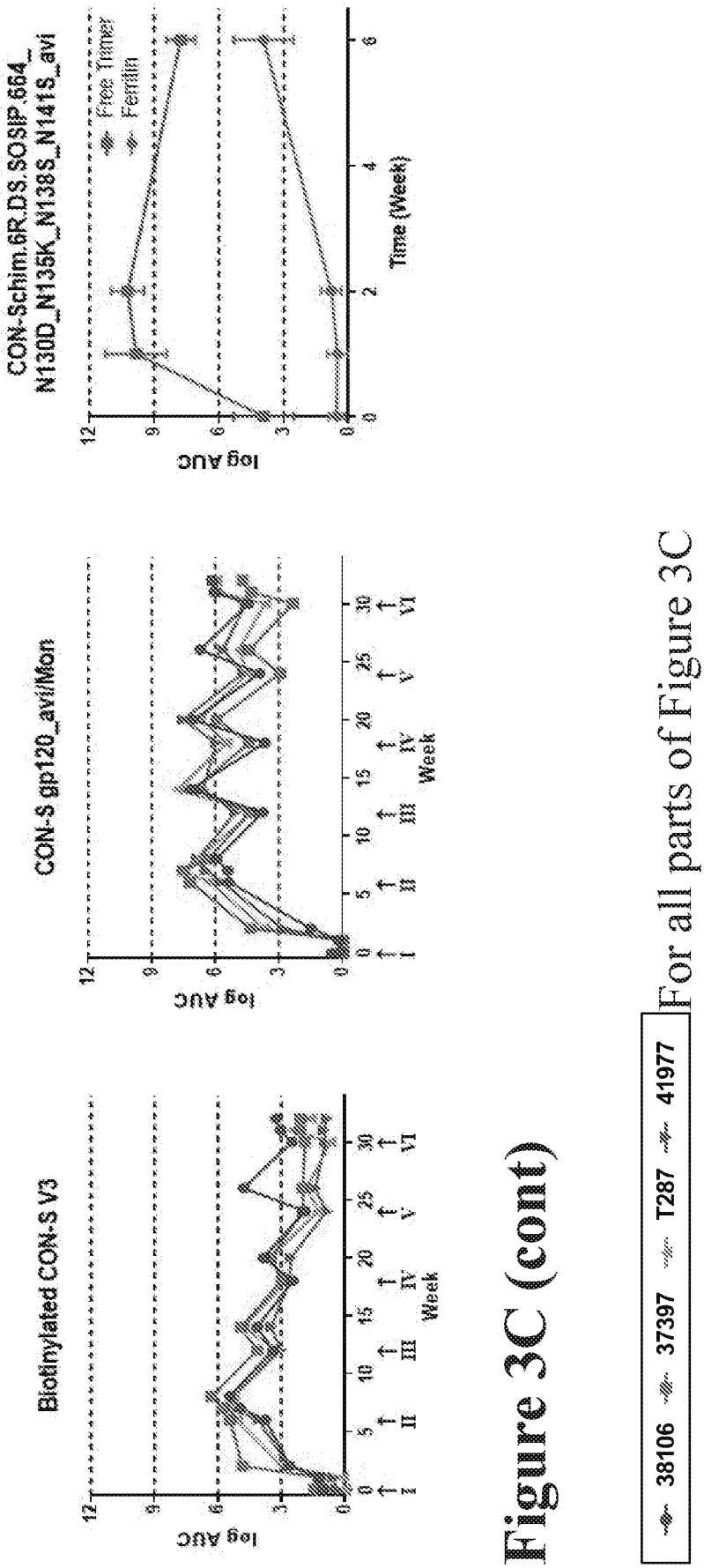


Figure 3C (cont)

For all parts of Figure 3C

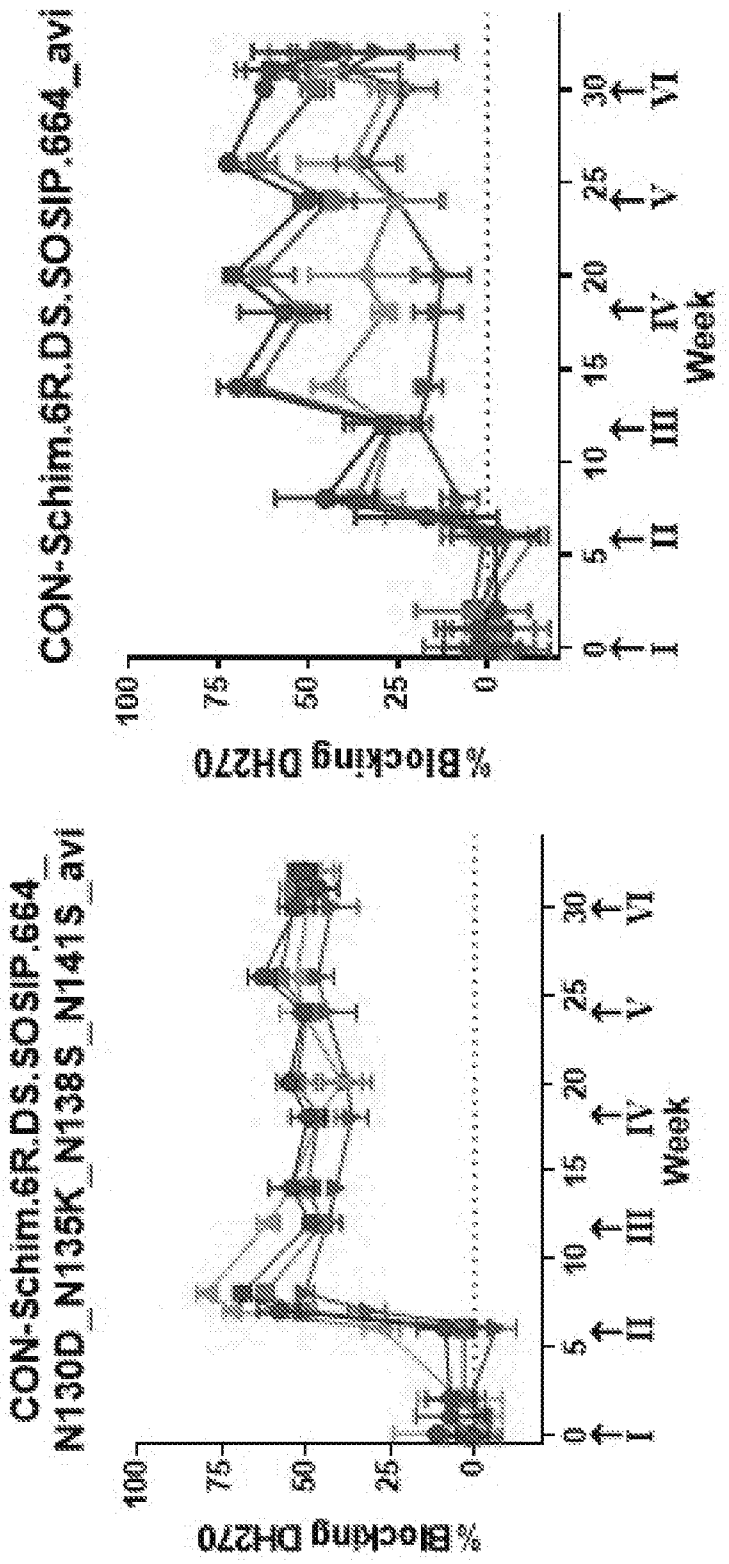

CON-Schim.6R.DS.SOSIP.664_avi
% Blocking DH270
100
75
50
25
0
0
5
10
15
20
25
30
I
II
III
IV
V
VI
Week
CON-Schim.6R.DS.SOSIP.664_
N130D_N135K_N138S_N141S_avi
% Blocking DH270
100
75
50
25
0
0
5
10
15
20
25
30
I
II
III
IV
V
VI
Week

Figure 4

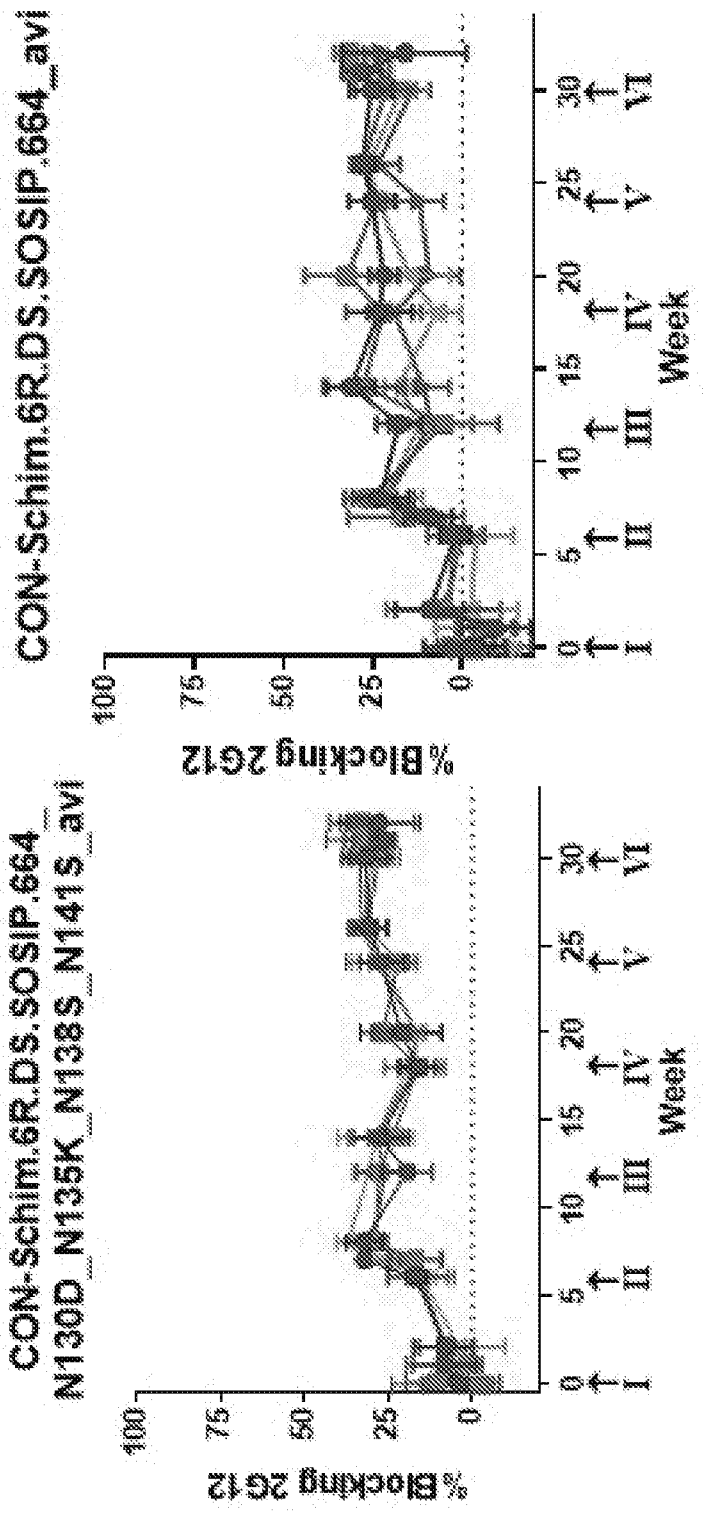


Figure 4 (cont)

For all parts of the figure



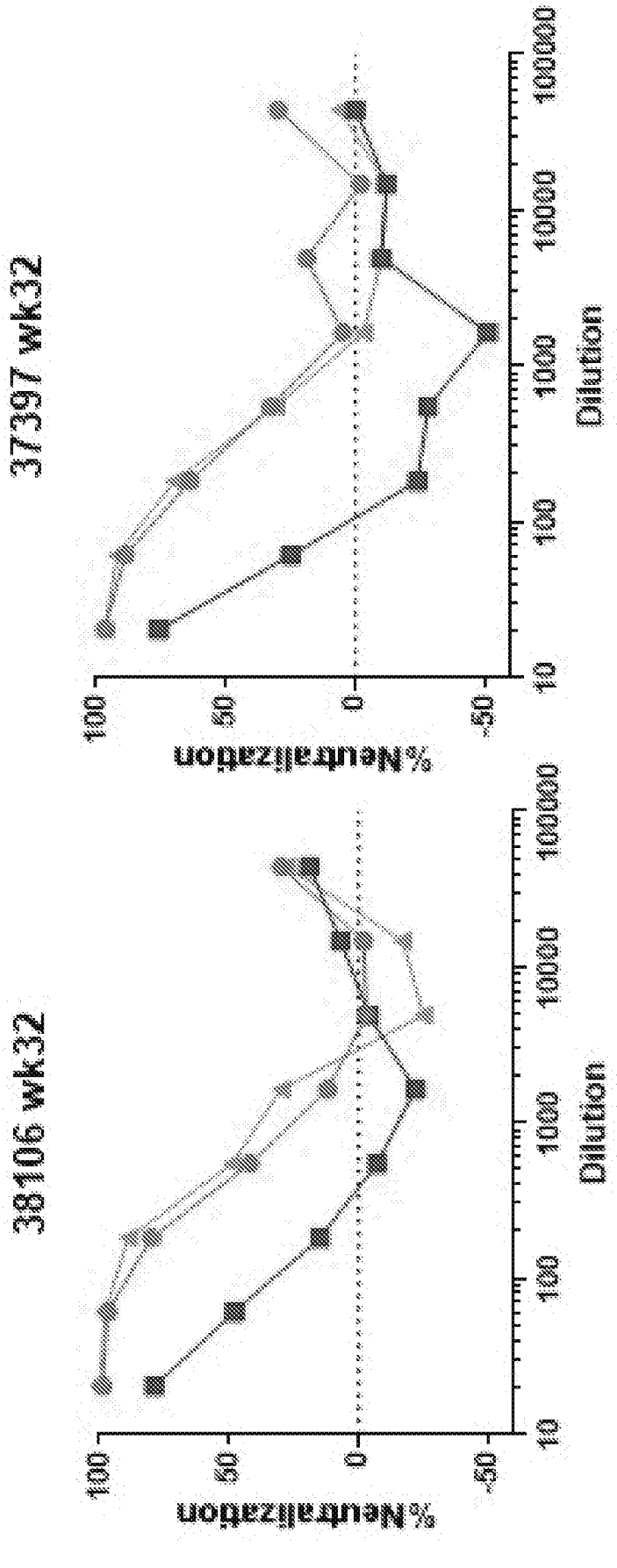

37397 wk32
% Neutralization
100
50
0
-50
10
100
1000
10000
100000
Dilution
38106 wk32
% Neutralization
100
50
0
-50
10
100
1000
10000
100000
Dilution

Figure 5

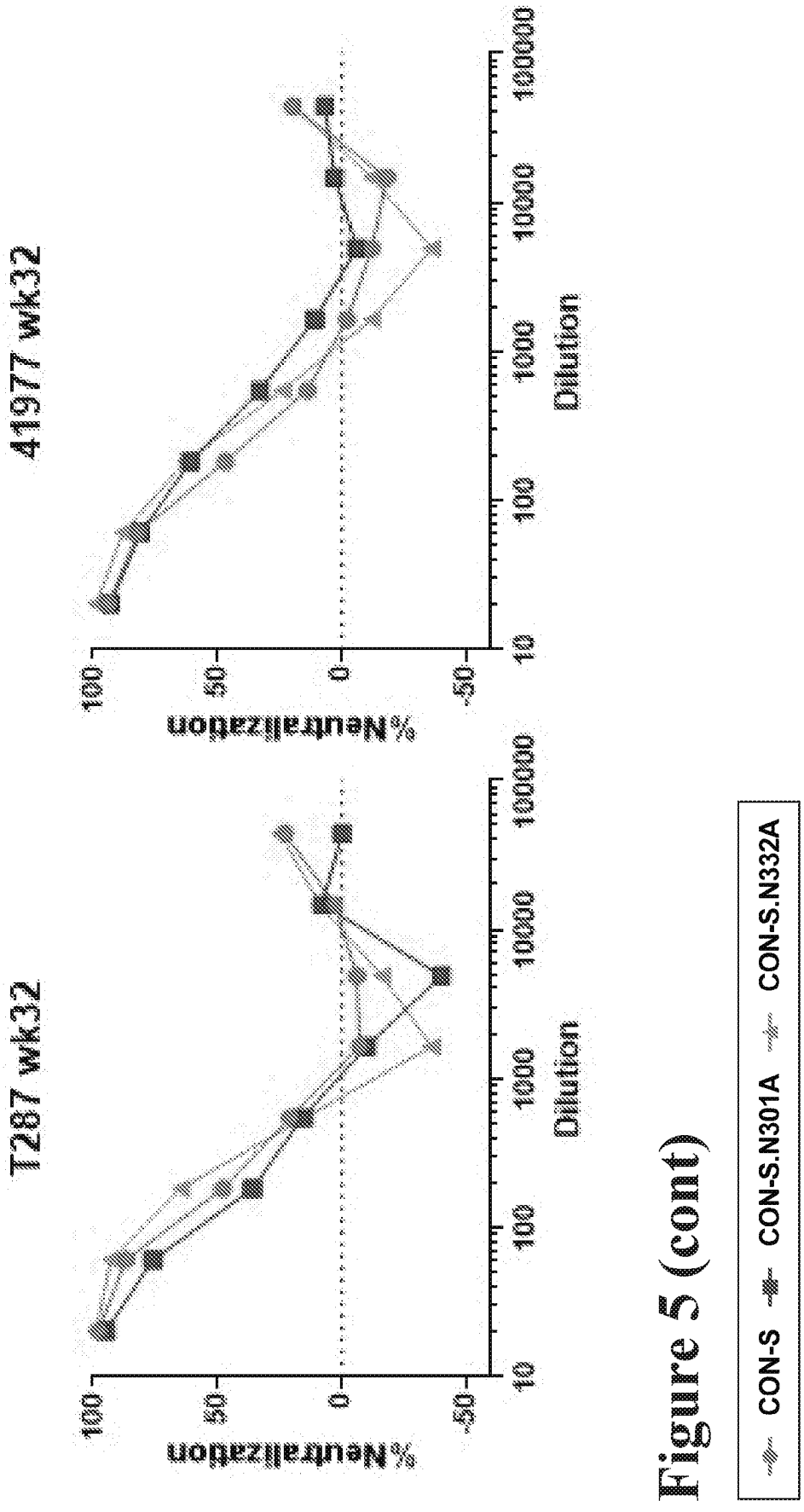


Figure 5 (cont)

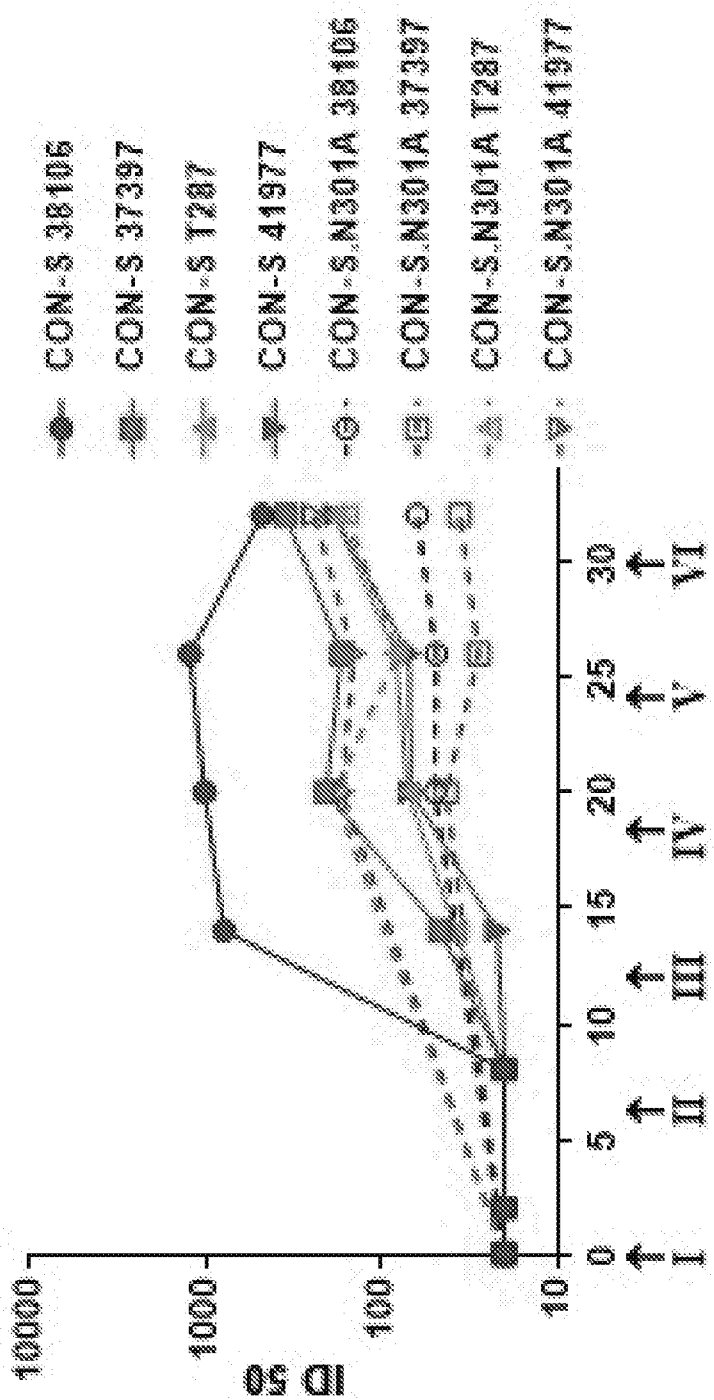


Figure 5 (cont)

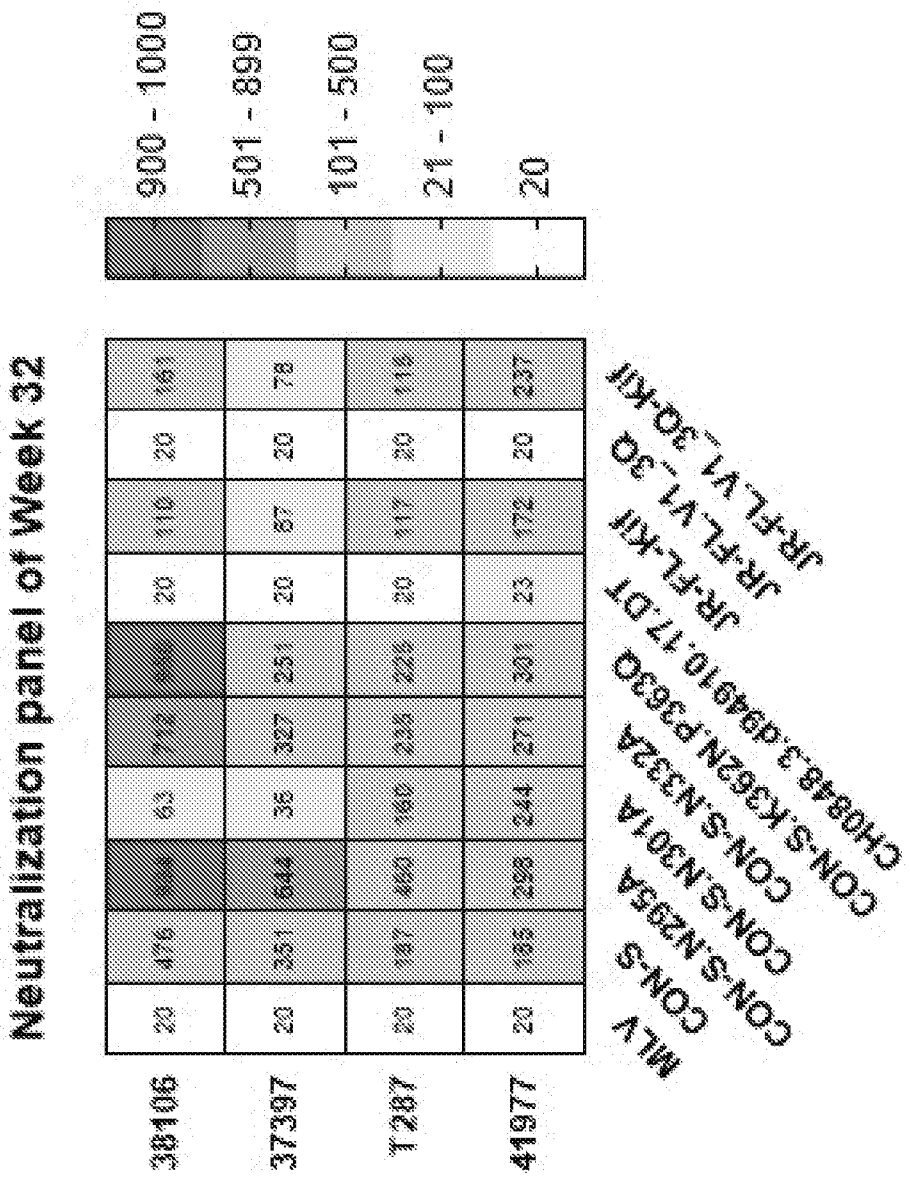


Figure 5 (cont)

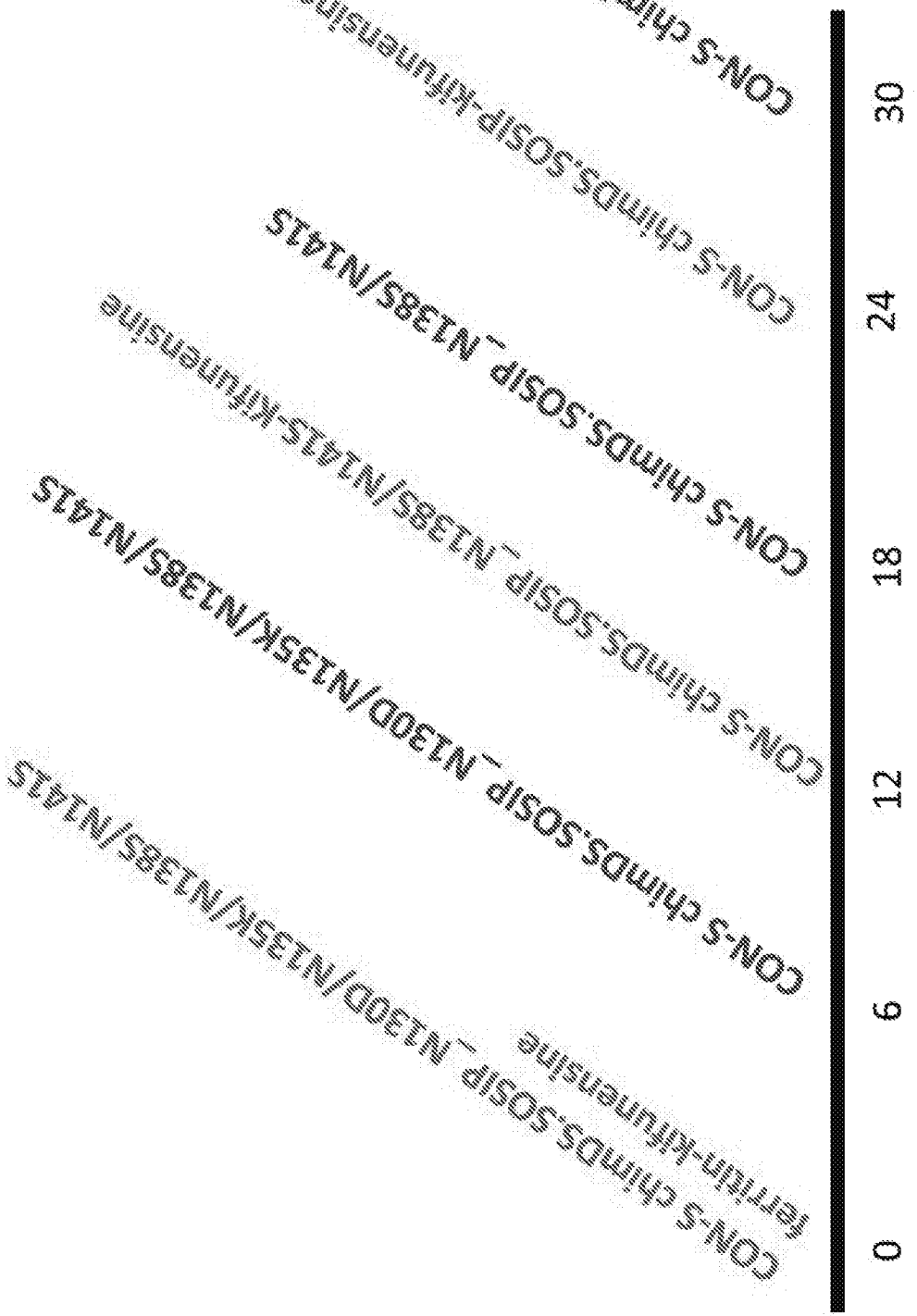

CON-S chimDS-SOSIP_N130D/N135K/N138S/N142S
CON-S chimDS-SOSIP_N130D/N135K/N138S/N142S
CON-S chimDS-SOSIP_N138S/N142S-Kifunensine
CON-S chimDS-SOSIP_N138S/N142S
CON-S chimDS-SOSIP-Kifunensine
CON-S chimDS-SOSIP
Ferritin-Kifunensine
0
6
12
18
24
30

Figure 6

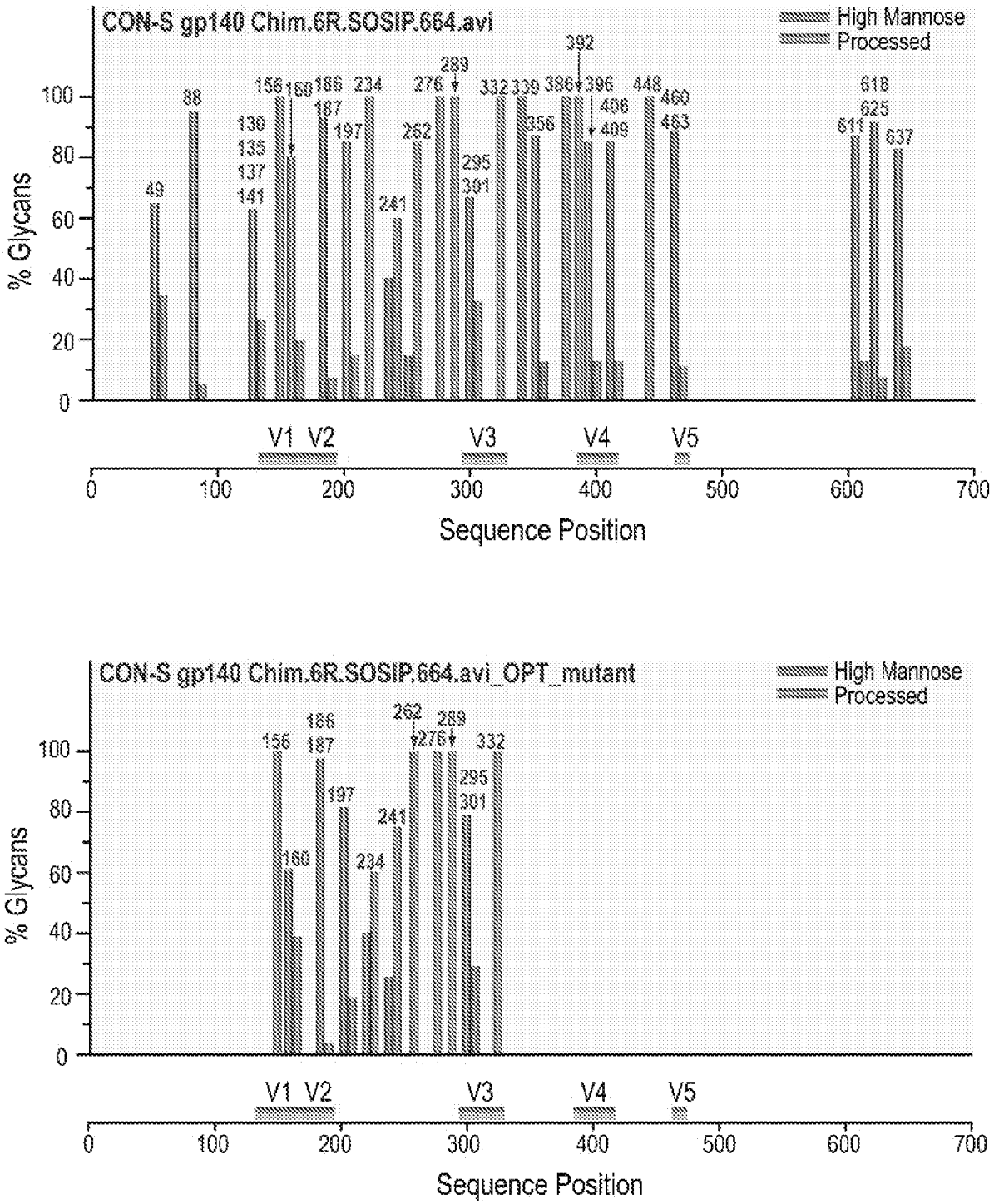

CON-S gp140 Chim.6R.SOSIP.664.avi
High Mannose
Processed
% Glycans
Sequence Position
V1 V2
V3
V4
V5
CON-S gp140 Chim.6R.SOSIP.664.avi_OPT_mutant
High Mannose
Processed
% Glycans
Sequence Position
V1 V2
V3
V4
V5

Figure 7

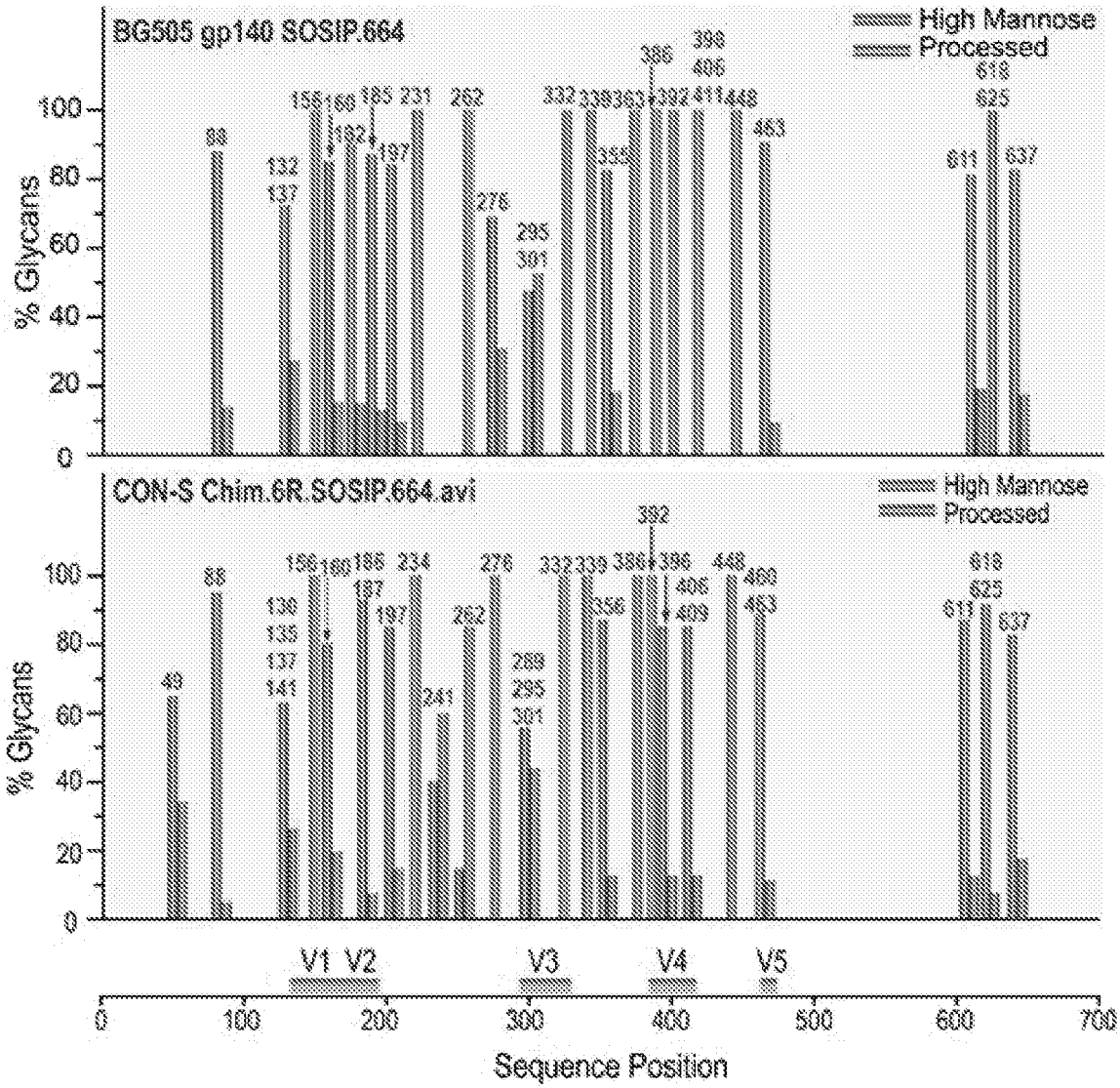

BG505 gp140 SOSIP.664
High Mannose
Processed
% Glycans
CON-S Chim.6R.SOSIP.664.avi
V1 V2
V3
V4
V5
Sequence Position

Figure 8

Disulfide Loop Domain	Disulfide-Linked Peptides	CS	Theoretical *m/z*	Experimental *m/z*	Mass Error (ppm)
I	EADTTLFC[carbamidomethyl]ASDAK AYDTEVHNVWATHAC[carbamidomethyl]VPTDPNPQEIVLEC[carbamidomethyl]VTENFNMWK	4+	1446.1578	1446.1647	5
		5+	1157.1277	1157.1355	7
		6+	964.4410	964.4479	7
		7+	826.8076	826.8113	5
II	QIINMWQGVGQC[carbamidomethyl]MYAPPIEGK NNMVEQMHEDIISLWDQSLKPC[carbamidomethyl]VK LINC[carbamidomethyl]D[deamidated]TSAC[carbamidomethyl]TQAC[carbamidomethyl]PK LTPLC[carbamidomethyl]VTLDC[carbamidomethyl]TNVK D[deamidated]C[carbamidomethyl]SFD[deamidated]ITTEIR	8+	1200.3159	1200.3250	8
		9+	1067.0594	1067.0637	4
		10+	960.4542	960.4531	1
III	C[carbamidomethyl]N[deamidated]KKFD[deamidated]GTGPC[carbamidomethyl]K	2+	705.8105	705.8138	5
		3+	470.8761	470.8779	4
	D[deamidated]VSTVQC[carbamidomethyl]THGIKPVVSTQLLLD[deamidated]GSLAEEEIIIR VSFEPIPIHYC[carbamidomethyl]APAGFAILK	4+	1434.2623	1434.2704	6
		5+	1147.6113	1147.6192	7
		6+	956.5106	956.5169	7
		7+	820.0101	820.0149	6

Figure 9A

SUBSTITUTE SHEET (RULE 26)

IV	TIIVQLDCESVEIDCCTRPNDCCNTR	3+	1196.2457	1196.2523	6
		4+	897.4361	897.4413	6
	QAHCCCCISGTK	5+	718.1503	718.1542	5
V	TIIFKPSSGGDLEITTHSFNCCR	3+	962.1492	962.1545	6
		4+	721.8637	721.8677	6
	ITCCK	5+	577.6924	577.6958	6
	GEFFYCCCTSGLFDCCSTWIGDCCGTK	3+	1344.9106	1344.9186	6
		4+	1008.9348	1008.9432	8
	NNDCCNTDCCDTITLPCCCR	5+	807.3493	807.3538	6
gp41 SOSIP	CCCK	3+	1177.2088	1177.2169	7
		4+	883.1584	883.1654	8
	LICCCCTNVPWDCCSSWSNR	5+	706.7282	706.7333	7
	DQQLLGIWGCCCSGK	6+	589.1080	589.1124	7

Figure 9A (cont)

Disulfide Loop Domain	Disulfide-Linked Peptides	CS	Theoretical *m/z*	Experimental *m/z*	Mass Error (ppm)
II	LTPLC*VTLDC*TNVK	2+	759.3888	759.3931	6
		3+	506.5949	506.5976	5
gp41	LIC*C*TNVPW[illegible]SSWSNR	2+	939.9167	939.9222	6
SOSIP		3+	626.9469	626.9505	6

Figure 9B

>HV1301184
GTCGACAAGCTTGCCACCATGAGGGTCCGGGGAATCCAGCGCAACTGCCAGCACCTCTGGAGGTGGGGCACGCTGAT
CCTGGGGATGCTGATGATCTGCAGCGCGGCTGAGAACCTGTGGGTGACAGTGTACTACGGCGTGCCTGTGTGGAAGG
AGGCCAACACCACCCTGTTCTGCGCCTCGGACGCCAAGGCCTACGACACGGAGGTCCACAACGTGTGGGCTACCCAC
GCCTGCGTGCCCACCGACCCCAATCCTCAGGAGATCGTCCTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAA
CAACATGGTGGAGCAGATGCACGAGGACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCC
CCCTGTGCGTGACCCTGAACTGCACGAACGTGAACGTGACCAACACCACGAACAACACGGAGGAGAAGGGGGAGATC
AAGAACTGCAGCTTCAACATCACCACCGAGATCCGGGACAAGAAGCAGAAGGTGTACGCCCTGTTCTACCGGCTGGA
CGTCGTGCCGATCGACGACAACAACAACAACTCCAGCAACTACAGGCTGATCAACTGCAACACCAGCGCGATCACCC
AGGCCTGCCCTAAGGTGTCGTTCGAGCCCATCCCCATCCACTACTGCGCGCCTGCCGGCTTCGCCATCCTGAAGTGC
AACGACAAGAAGTTCAACGGCACCGGCCCCTGCAAGAACGTCAGCACCGTCCAGTGCACCCACGGCATCAAGCCTGT
GGTGTCCACCCAGCTGCTCCTGAACGGCAGCCTGGCCGAGGAGGAGATCATCATCAGGAGCGAGAACATCACCAACA
ACGCCAAGACGATCATCGTGCAGCTGAACGAGTCGGTGGAGATCAACTGCACCCGGCCCAACAACAACACGCGGAAG
AGCATCCGGATCGGCCCTGGACAGGCGTTCTACGCCACGGGCGACATCATCGGCGACATCAGGCAGGCCCACTGCAA
CATCTCGGGGACGAAGTGGAACAAGACCCTGCAGCAGGTCGCGAAGAAGCTGAGGGAGCACTTCAACAACAAGACCA
TCATCTTCAAGCCGAGCAGCGGCGGAGACCTGGAGATCACCACGCACTCGTTCAACTGCCGGGGCGAGTTCTTCTAC
TGTAACACGTCGGGCCTGTTCAACAGCACCTGGATCGGCAACGGCACGAAGAACAACAACAACACTAACGACACCAT
CACCCTGCCCTGCCGGATCAAGCAGATCATCAACATGTGGCAGGGCGTGGGCCAGGCTATGTACGCCCCTCCCATCG
AGGGCAAGATCACGTGCAAGAGCAACATCACCGGCCTGCTGCTGACCAGGGACGGCGGGAACAACAACACGAACGAG
ACCGAGATCTTCAGACCTGGCGGCGGAGACATGAGAGACAACTGGCGGAGCGAGCTGTACAAGTACAAGGTCGTGAA
GATCGAGCCCCTGGGCGTCGCACCCACCAAGTGCAAGGAGAGGGTGGTGGGCAGGCGACGCCGTAGGCGGGCGGTCG
GCATCGGCGCCGTGTTCCTGGGCTTCCTGGGAGCAGCCGGCAGCACCATGGGAGCCGCCTCGATCACCCTGACCGTG
CAGGCGAGGCAGCTGCTGTCCGGCATCGTGCAGCAGCAGTCGAACCTGCTGAGGGCCCCCGAGGCCCAGCAGCACCT
GCTCCAGCTGACCGTGTGGGGCATCAAGCAGCTCCAGGCCAGGGTGCTGGCCGTCGAGCGCTACCTGAAGGACCAGC
AGCTGCTCGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCTGCACCACCGTGCCCTGGAACAGCAGCTGGAGCAAC
AAGAGCCAGGACGAGATCTGGGACAACATGACCTGGATGGAGTGGGAGCGGGAGATCAACAACTACACCGACATCAT
CTACAGCCTGATCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGCGCTGGACTGATCTAGAG
GATCC

>HV1301185
GTCGACAAGCTTGCCACCATGAGGGTCCGGGGAATCCAGCGCAACTGCCAGCACCTCTGGAGGTGGGGCACGCTGAT
CCTGGGGATGCTGATGATCTGCAGCGCGGCTGAGAACCTGTGGGTGACAGTGTACTACGGCGTGCCTGTGTGGAAGG
AGGCCAACACCACCCTGTTCTGCGCCTCGGACGCCAAGGCCTACGACACGGAGGTCCACAACGTGTGGGCTACCCAC
GCCTGCGTGCCCACCGACCCCAATCCTCAGGAGATCGTCCTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAA
CAACATGGTGGAGCAGATGCACGAGGACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCC
CCCTGTGCGTGACCCTGAACTGCACGAACGTGAACGTGACCAACACCACGAACAACACGGAGGAGAAGGGGGAGATC
AAGAACTGCAGCTTCAACATCACCACCGAGATCCGGGACAAGAAGCAGAAGGTGTACGCCCTGTTCTACCGGCTGGA
CGTCGTGCCGATCGACGACAACAACAACAACTCCAGCAACTACAGGCTGATCAACTGCAACACCAGCGCGWGCACCC
AGGCCTGCCCTAAGGTGTCGTTCGAGCCCATCCCCATCCACTACTGCGCGCCTGCCGGCTTCGCCATCCTGAAGTGC
AACGACAAGAAGTTCAACGGCACCGGCCCCTGCAAGAACGTCAGCACCGTCCAGTGCACCCACGGCATCAAGCCTGT
GGTGTCCACCCAGCTGCTCCTGAACGGCAGCCTGGCCGAGGAGGAGATCATCATCAGGAGCGAGAACATCACCAACA
ACGCCAAGACGATCATCGTGCAGCTGAACGAGTCGGTGGAGATCAACTGCACCCGGCCCAACAACAACACGCGGAAG
AGCATCCGGATCGGCCCTGGACAGGCGTTCTACGCCACGGGCGACATCATCGGCGACATCAGGCAGGCCCACTGCAA
CATCTCGGGGACGAAGTGGAACAAGACCCTGCAGCAGGTCGCGAAGAAGCTGAGGGAGCACTTCAACAACAAGACCA
TCATCTTCAAGCCGAGCAGCGGCGGAGACCTGGAGATCACCACGCACTCGTTCAACTGCCGGGGCGAGTTCTTCTAC
TGTAACACGTCGGGCCTGTTCAACAGCACCTGGATCGGCAACGGCACGAAGAACAACAACAACACTAACGACACCAT
CACCCTGCCCTGCCGGATCAAGCAGATCATCAACATGTGGCAGGGCGTGGGCCAGWGTATGTACGCCCCTCCCATCG
AGGGCAAGATCACGTGCAAGAGCAACATCACCGGCCTGCTGCTGACCAGGGACGGCGGGAACAACAACACGAACGAG
ACCGAGATCTTCAGACCTGGCGGCGGAGACATGAGAGACAACTGGCGGAGCGAGCTGTACAAGTACAAGGTCGTGAA
GATCGAGCCCCTGGGCGTCGCACCCACCAAGTGCAAGGAGAGGGTGGTGGGCAGGCGACGCCGTAGGCGGGCGGTCG
GCATCGGCGCCGTGTTCCTGGGCTTCCTGGGAGCAGCCGGCAGCACCATGGGAGCCGCCTCGATCACCCTGACCGTG

Figure 10

CAGGCGAGGCAGCTGCTGTCCGGCATCGTGCAGCAGCAGTCGAACCTGCTGAGGGCCCCCGAGGCCCAGCAGCACCT
GCTCCAGCTGACCGTGTGGGGCATCAAGCAGCTCCAGGCCAGGGTGCTGGCCGTCGAGCGCTACCTGAAGGACCAGC
AGCTGCTCGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCTGCACCACCGTGCCCTGGAACAGCAGCTGGAGCAAC
AAGAGCCAGGACGAGATCTGGGACAACATGACCTGGATGGAGTGGGAGCGGGAGATCAACAACTACACCGACATCAT
CTACAGCCTGATCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGCGCTGGACTGATCTAGAG
GATCC

>HV1301186
GTCGACAAGCTTGCCACCATGAGGGTCCGGGGAATCCAGCGCAACTGCCAGCACCTCTGGAGGTGGGGCACGCTGAT
CCTGGGGATGCTGATGATCTGCAGCGCGGCTGAGAACCTGTGGGTGACAGTGTACTACGGCGTGCCTGTGTGGAAGG
AGGCCAACACCACCCTGTTCTGCGCCTCGGACGCCAAGGCCTACGACACGGAGGTCCACAACGTGTGGGCTACCCAC
GCCTGCGTGCCCACCGACCCCAATCCTCAGGAGATCGTCCTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAA
CAACATGGTGGAGCAGATGCACGAGGACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCC
CCCTGTGCGTGACCCTGAACTGCACGAACGTGAACGTGACCAACACCACGAACAACACGGAGGAGAAGGGGGAGATC
AAGAACTGCAGCTTCAACATCACCACCGAGATCCGGGACAAGAAGCAGAAGGTGTACGCCCTGTTCTACCGGCTGGA
CGTCGTGCCGATCGACGACAACAACAACAACTCCAGCAACTACAGGCTGATCAACTGCAACACCAGCGCGATCACCC
AGGCCTGCCCTAAGGTGTCGTTCGAGCCCATCCCCATCCACTACTGCGCGCCTGCCGGCTTCGCCATCCTGAAGTGC
AACGACAAGAAGTTCAACGGCACCGGCCCCTGCAAGAACGTCAGCACCGTCCAGTGCACCCACGGCATCAAGCCTGT
GGTGTCCACCCAGCTGCTCCTGAACGGCAGCCTGGCCGAGGAGGAGATCATCATCAGGAGCGAGAACATCACCAACA
ACGCCAAGACGATCATCGTGCAGCTGAACGAGTCGGTGGAGATCAACTGCACCCGGCCCAACAACAACACGCGGAAG
AGCATCCGGATCGGCCCTGGACAGGCGTTCTACGCCACGGGCGACATCATCGGCGACATCAGGCAGGCCCACTGCAA
CATCTCGGGGACGAAGTGGAACAAGACCCTGCAGCAGGTCGCGAAGAAGCTGAGGGAGCACTTCAACAACAAGACCA
TCATCTTCAAGCCGAGCAGCGGCGGAGACCTGGAGATCACCACGCACTCGTTCAACTGCCGGGGCGAGTTCTTCTAC
TGTAACACGTCGGGCCTGTTCAACAGCACCTGGATCGGCAACGGCACGAAGAACAACAACAACACTAACGACACCAT
CACCCTGCCCTGCCGGATCAAGCAGATCATCAACATGTGGCAGGGCGTGGGCCAGGCTATGTACGCCCCTCCCATCG
AGGGCAAGATCACGTGCAAGAGCAACATCACCGGCCTGCTGCTGACCAGGGACGGCGGGAACAACAACACGAACGAG
ACCGAGATCTTCAGACCTGGCGGCGGAGACATGAGAGACAACTGGCGGAGCGAGCTGTACAAGTACAAGGTCGTGAA
GATCGAGCCCCTGGGCGTCGCACCCACCAAGTGCAAGGAGAGGGTGGTGGGCAGGCGACGCCGTAGGCGGGCGGTCG
GCATCGGCGCCGTGTTCCTGGGCTTCCTGGGAGCAGCCGGCAGCACCATGGGAGCCGCCTCGATGACCCTGACCGTG
CAGGCGAGGCAGCTGCTGTCCGGCATCGTGCAGCAGCAGTCGAACCTGCTGAGGGCCCCCGAGGCCCAGCAGCACCT
GCTCCAGCTGACCGTGTGGGGCATCAAGCAGCTCCAGGCCAGGGTGCTGGCCGTCGAGCGCTACCTGAAGGACCAGC
AGCTGCTCGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCTGCACCACCGTGCCCTGGAACAGCAGCTGGAGCAAC
AAGAGCCAGGACGAGATCTGGGACAACATGACCTGGATGGAGTGGGAGCGGGAGATCAACAACTACACCGACATCAT
CTACAGCCTGATCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGCGCTGGACTGATCTAGAG
GATCC

>HV1301187
GTCGACAAGCTTGCCACCATGAGGGTCCGGGGAATCCAGCGCAACTGCCAGCACCTCTGGAGGTGGGGCACGCTGAT
CCTGGGGATGCTGATGATCTGCAGCGCGGCTGAGAACCTGTGGGTGACAGTGTACTACGGCGTGCCTGTGTGGAAGG
AGGCCAACACCACCCTGTTCTGCGCCTCGGACGCCAAGGCCTACGACACGAAGGTCCACAACGTGTGGGCTACCCAC
GCCTGCGTGCCCACCGACCCCAATCCTCAGGAGATCGTCCTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAA
CAACATGGTGGAGCAGATGCACGAGGACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCC
CCCTGTGCGTGACCCTGAACTGCACGAACGTGAACGTGACCAACACCACGAACAACACGGAGGAGAAGGGGGAGATC
AAGAACTGCAGCTTCAACATCACCACCGAGATCCGGGACAAGAAGCAGAAGGTGTACGCCCTGTTCTACCGGCTGGA
CGTCGTGCCGATCGACGACAACAACAACAACTCCAGCAACTACAGGCTGATCAACTGCAACACCAGCGCGATCACCC
AGGCCTGCCCTAAGGTGTCGTTCGAGCCCATCCCCATCCACTACTGCGCGCCTGCCGGCTTCGCCATCCTGAAGTGC
AACGACAAGAAGTTCAACGGCACCGGCCCCTGCAAGAACGTCAGCACCGTCCAGTGCACCCACGGCATCAAGCCTGT
GGTGTCCACCCAGCTGCTCCTGAACGGCAGCCTGGCCGAGGAGGAGATCATCATCAGGAGCGAGAACATCACCAACA
ACGCCAAGACGATCATCGTGCAGCTGAACGAGTCGGTGGAGATCAACTGCACCCGGCCCAACAACAACACGCGGAAG
AGCATCCGGATCGGCCCTGGACAGTGGTTCTACGCCACGGGCGACATCATCGGCGACATCAGGCAGGCCCACTGCAA
CATCTCGGGGACGAAGTGGAACAAGACCCTGCAGCAGGTCGCGAAGAAGCTGAGGGAGCACTTCAACAACAAGACCA
TCATCTTCAAGCCGAGCAGCGGCGGAGACCTGGAGATCACCACGCACTCGTTCAACTGCCGGGGCGAGTTCTTCTAC
TGTAACACGTCGGGCCTGTTCAACAGCACCTGGATCGGCAACGGCACGAAGAACAACAACAACACTAACGACACCAT
CACCCTGCCCTGCCGGATCAAGCAGATCATCAACATGTGGCAGGGCGTGGGCCAGGCTATGTACGCCCCTCCCATCG

Figure 10 (cont)

AGGGCAAGATCACGTGCAAGAGCAACATCACCGGCCTGCTGCTGACCAGGGACGGCGGGAACAACAACACGAACGAG
ACCGAGATCTTCAGACCTGGCGGCGGAGACATGAGAGACAACTGGCGGAGCGAGCTGTACAAGTACAAGGTCGTGAA
GATCGAGCCCCTGGGCGTCGCACCCACCAAGTGCAAGGAGAGGGTGGTGGGCAGGCGACGCCGTAGGCGGGCGGTCG
GCATCGGCGCCGTGTTCCTGGGCTTCCTGGGAGCAGCCGGCAGCACCATGGGAGCCGCCTCGATGACCCTGACCGTG
CAGGCGAGGCAGCTGCTGTCCGGCATCGTGCAGCAGCAGTCGAACCTGCTGAGGGCCCCCGAGGCCCAGCAGCACCT
GCTCCAGCTGACCGTGTGGGGCATCAAGCAGCTCCAGGCCAGGGTGCTGGCCGTCGAGCGCTACCTGAAGGACCAGC
AGCTGCTCGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCTGCACCACCGTGCCCTGGAACAGCAGCTGGAGCAAC
AAGAGCCAGGACGAGATCTGGGACAACATGACCTGGATGGAGTGGGAGCGGGAGATCAACAACTACACCGACATCAT
CTACAGCCTGATCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGCGCTGGACTGATCTAGAG
GATCC

>HV1301188
GTCGACAAGCTTGCCACCATGAGGGTCCGGGGAATCCAGCGCAACTGCCAGCACCTCTGGAGGTGGGGCACGCTGAT
CCTGGGGATGCTGATGATCTGCAGCGCGGCTGAGAACCTGTGGGTGACAGTGTACTACGGCGTGCCTGTGTGGAAGG
AGGCCAACACCACCCTGTTCTGCGCCTCGGACGCCAAGGCCTACGACACGGAGGTCAGAAACGTGTGGGCTACCCAC
GCCTGCGTGCCCACCGACCCCAATCCTCAGGAGATCGTCCTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAA
CAACATGGTGGAGCAGATGCACGAGGACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCC
CCCTGTGCGTGACCCTGAACTGCACGAACGTGAACGTGACCAACACCACGAACAACACGGAGGAGAAGGGGGAGATC
AAGAACTGCAGCTTCAACATCACCACCGAGATCCGGGACAAGAAGCAGAAGGTGTACGCCCTGTTCTACCGGCTGGA
CGTCGTGCCGATCGACGACAACAACAACAACTCCAGCAACTACAGGCTGATCAACTGCAACACCAGCGCGATCACCC
AGGCCTGCCCTAAGGTGTCGTTCGAGCCCATCCCCATCCACTACTGCGCGCCTGCCGGCTTCGCCATCCTGAAGTGC
AACGACAAGAAGTTCAACGGCACCGGCCCCTGCAAGAACGTCAGCACCGTCCAGTGCACCCACGGCATCAAGCCTGT
GGTGTCCACCCAGCTGCTCCTGAACGGCAGCCTGGCCGAGGAGGAGATCATCATCAGGAGCGAGAACATCACCAACA
ACGCCAAGACGATCATCGTGCAGCTGAACGAGTCGGTGGAGATCAACTGCACCCGGCCCAACAACAACACGCGGAAG
AGCATCCGGATCGGCCCTGGACAGTGGTTCTACGCCACGGGCGACATCATCGGCGACATCAGGCAGGCCCACTGCAA
CATCTCGGGGACGAAGTGGAACAAGACCCTGCAGCAGGTCGCGAAGAAGCTGAGGGAGCACTTCAACAACAAGACCA
TCATCTTCAAGCCGAGCAGCGGCGGAGACCTGGAGATCACCACGCACTCGTTCAACTGCCGGGGCGAGTTCTTCTAC
TGTAACACGTCGGGCCTGTTCAACAGCACCTGGATCGGCAACGGCACGAAGAACAACAACAACACTAACGACACCAT
CACCCTGCCCTGCCGGATCAAGCAGATCATCAACATGTGGCAGGGCGTGGGCCAGGCTATGTACGCCCCTCCCATCG
AGGGCAAGATCACGTGCAAGAGCAACATCACCGGCCTGCTGCTGACCAGGGACGGCGGGAACAACAACACGAACGAG
ACCGAGATCTTCAGACCTGGCGGCGGAGACATGAGAGACAACTGGCGGAGCGAGCTGTACAAGTACAAGGTCGTGAA
GATCGAGCCCCTGGGCGTCGCACCCACCAAGTGCAAGGAGAGGGTGGTGGGCAGGCGACGCCGTAGGCGGGCGGTCG
GCATCGGCGCCGTGTTCCTGGGCTTCCTGGGAGCAGCCGGCAGCACCATGGGAGCCGCCTCGATGACCCTGACCGTG
CAGGCGAGGCAGCTGCTGTCCGGCATCGTGCAGCAGCAGTCGAACCTGCTGAGGGCCCCCGAGGCCCAGCAGCACCT
GCTCCAGCTGACCGTGTGGGGCATCAAGCAGCTCCAGGCCAGGGTGCTGGCCGTCGAGCGCTACCTGAAGGACCAGC
AGCTGCTCGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCTGCACCACCGTGCCCTGGAACAGCAGCTGGAGCAAC
AAGAGCCAGGACGAGATCTGGGACAACATGACCTGGATGGAGTGGGAGCGGGAGATCAACAACTACACCGACATCAT
CTACAGCCTGATCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGCGCTGGACTGATCTAGAG
GATCC
>HV1301257
GTCGACGCCACCATGGGCTCCCTGCAGCCCCTGGCCACCCTGTACCTGCTGGGCATGCTGGTGGCCTCCGTGCTGGC
CGCCGAGAACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCCAACACCACCCTGTTCTGCGCCT
CCGACGCCAAGGCCTACGACACCGAGGTGCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCC
CAGGAGATCGTGCTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAGGA
CATCATCTCCCTGTGGGACCAGTCCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTGACCCTGAACTGCACCA
ACGTGAACGTGACCAACACCACCAACAACACCGAGGAGAAGGGCGAGATCAAGAACTGCTCCTTCAACATCACCACC
GAGATCCGCGACAAGAAGCAGAAGGTGTACGCCCTGTTCTACCGCCTGGACGTGGTGCCCATCGACGACAACAACAA
CAACTCCTCCAACTACCGCCTGATCAACTGCAACACCTCCGCCATCACCCAGGCCTGCCCCAAGGTGTCCTTCGAGC
CCATCCCCATCCACTACTGCGCCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCACCGGC
CCCTGCAAGAACGTGTCCACCGTGCAGTGCACCCACGGCATCAAGCCCGTGGTGTCCACCCAGCTGCTGCTGAACGG
CTCCCTGGCCGAGGAGGAGATCATCATCCGCTCCGAGAACATCACCAACAACGCCAAGACCATCATCGTGCAGCTGA
ACGAGTCCGTGGAGATCAACTGCACCCGCCCCAACAACAACACCCGCAAGTCCATCCGCATCGGCCCCGGCCAGGCC
TTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCTCCGGCACCAAGTGGAACAAGAC
CCTGCAGCAGGTGGCCAAGAAGCTGCGCGAGCACTTCAACAACAAGACCATCATCTTCAAGCCCTCCTCCGGCGGCG

Figure 10 (cont)

ACCTGGAGATCACCACCCACTCCTTCAACTGCCGCGGCGAGTTCTTCTACTGCAACACCTCCGGCCTGTTCAACTCC
ACCTGGATCGGCAACGGCACCAAGAACAACAACAACACCAACGACACCATCACCCTGCCCTGCCGCATCAAGCAGAT
CATCAACATGTGGCAGGGCGTGGGCCAGGCCATGTACGCCCCCCCCATCGAGGGCAAGATCACCTGCAAGTCCAACA
TCACCGGCCTGCTGCTGACCCGCGACGGCGGCAACAACAACACCAACGAGACCGAGATCTTCCGCCCCGGCGGCGGC
GACATGCGCGACAACTGGCGCTCCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCAC
CCGCTGCAAGCGCCGCGTGGTGGGCCGCCGCCGCCGCCGCCGCGCCGTGGGCATCGGCGCCGTGTTCCTGGGCTTCC
TGGGCGCCGCCGGCTCCACCATGGGCGCCGCCTCCATGACCCTGACCGTGCAGGCCCGCAACCTGCTGTCCGGCATC
GTGCAGCAGCAGTCCAACCTGCTGCGCGCCCCCGAGGCCCAGCAGCACCTGCTGAAGCTGACCGTGTGGGGCATCAA
GCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGCGCGACCAGCAGCTGCTGGGCATCTGGGGCTGCTCCG
GCAAGCTGATCTGCTGCACCAACGTGCCCTGGAACTCCTCCTGGTCCAACCGCAACCTGTCCGAGATCTGGGACAAC
ATGACCTGGCTGCAGTGGGACAAGGAGATCTCCAACTACACCCAGATCATCTACGGCCTGCTGGAGGAGTCCCAGAA
CCAGCAGGAGAAGAACGAGCAGGACCTGCTGGCCCTGGACGGCTCCGGCCTGAACGACATCTTCGAGGCCCAGAAGA
TCGAGTGGCACGAGTAGGGATCC

>HV1301258

gtcgaccgccaccATGGGCTCCCTGCAGCCCCTGGCCACCCTGTACCTGCTGGGCATGCTGGTGGCCTCCGTGCTGGC
CGCCGAGAACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCCAACACCACCCTGTTCTGCGCCT
CCGACGCCAAGGCCTACGACACCGAGGTGCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCC
CAGGAGATCGTGCTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAGGA
CATCATCTCCCTGTGGGACCAGTCCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTGACCCTGAACTGCACCA
ACGTGAACGTGACCAACACCACCAACAACACCGAGGAGAAGGGCGAGATCAAGAACTGCTCCTTCAACATCACCACC
GAGATCCGCGACAAGAAGCAGAAGGTGTACGCCCTGTTCTACCGCCTGGACGTGGTGCCCATCGACGACAACAACAA
CAACTCCTCCAACTACCGCCTGATCAACTGCAACACCTCCGCC**TG**CACCCAGGCCTGCCCCAAGGTGTCCTTCGAGC
CCATCCCCATCCACTACTGCGCCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCACCGGC
CCCTGCAAGAACGTGTCCACCGTGCAGTGCACCCACGGCATCAAGCCCGTGGTGTCCACCCAGCTGCTGCTGAACGG
CTCCCTGGCCGAGGAGGAGATCATCATCCGCTCCGAGAACATCACCAACAACGCCAAGACCATCATCGTGCAGCTGA
ACGAGTCCGTGGAGATCAACTGCACCCGCCCCAACAACAACACCCGCAAGTCCATCCGCATCGGCCCCGGCCAGGCC
TTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCTCCGGCACCAAGTGGAACAAGAC
CCTGCAGCAGGTGGCCAAGAAGCTGCGCGAGCACTTCAACAACAAGACCATCATCTTCAAGCCCTCCTCCGGCGGCG
ACCTGGAGATCACCACCCACTCCTTCAACTGCCGCGGCGAGTTCTTCTACTGCAACACCTCCGGCCTGTTCAACTCC
ACCTGGATCGGCAACGGCACCAAGAACAACAACAACACCAACGACACCATCACCCTGCCCTGCCGCATCAAGCAGAT
CATCAACATGTGGCAGGGCGTGGGCCAG**TG**CATGTACGCCCCCCCCATCGAGGGCAAGATCACCTGCAAGTCCAACA
TCACCGGCCTGCTGCTGACCCGCGACGGCGGCAACAACAACACCAACGAGACCGAGATCTTCCGCCCCGGCGGCGGC
GACATGCGCGACAACTGGCGCTCCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCAC
CCGCTGCAAGCGCCGCGTGGTGGGCCGCCGCCGCCGCCGCCGCGCCGTGGGCATCGGCGCCGTGTTCCTGGGCTTCC
TGGGCGCCGCCGGCTCCACCATGGGCGCCGCCTCCATGACCCTGACCGTGCAGGCCCGCAACCTGCTGTCCGGCATC
GTGCAGCAGCAGTCCAACCTGCTGCGCGCCCCCGAGGCCCAGCAGCACCTGCTGAAGCTGACCGTGTGGGGCATCAA
GCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGCGCGACCAGCAGCTGCTGGGCATCTGGGGCTGCTCCG
GCAAGCTGATCTGCTGCACCAACGTGCCCTGGAACTCCTCCTGGTCCAACCGCAACCTGTCCGAGATCTGGGACAAC
ATGACCTGGCTGCAGTGGGACAAGGAGATCTCCAACTACACCCAGATCATCTACGGCCTGCTGGAGGAGTCCCAGAA
CCAGCAGGAGAAGAACGAGCAGGACCTGCTGGCCCTGGACGGCTCCGGCCTGAACGACATCTTCGAGGCCCAGAAGA
TCGAGTGGCACGAGtagggatcc

>HV1301259

GTCGACGCCACCATGGGCTCCCTGCAGCCCCTGGCCACCCTGTACCTGCTGGGCATGCTGGTGGCCTCCGTGCTGGC
CGCCGAGAACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCCAACACCACCCTGTTCTGCGCCT
CCGACGCCAAGGCCTACGACACCAAGGTGCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCC
CAGGAGATCGTGCTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAGGA
CATCATCTCCCTGTGGGACCAGTCCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTGACCCTGAACTGCACCA
ACGTGAACGTGACCAACACCACCAACAACACCGAGGAGAAGGGCGAGATCAAGAACTGCTCCTTCAACATCACCACC
GAGATCCGCGACAAGAAGCAGAAGGTGTACGCCCTGTTCTACCGCCTGGACGTGGTGCCCATCGACGACAACAACAA
CAACTCCTCCAACTACCGCCTGATCAACTGCAACACCTCCGCCATCACCCAGGCCTGCCCCAAGGTGTCCTTCGAGC
CCATCCCCATCCACTACTGCGCCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCACCGGC
CCCTGCAAGAACGTGTCCACCGTGCAGTGCACCCACGGCATCAAGCCCGTGGTGTCCACCCAGCTGCTGCTGAACGG
CTCCCTGGCCGAGGAGGAGATCATCATCCGCTCCGAGAACATCACCAACAACGCCAAGACCATCATCGTGCAGCTGA
ACGAGTCCGTGGAGATCAACTGCACCCGCCCCAACAACAACACCCGCAAGTCCATCCGCATCGGCCCCGGCCAGTGG

Figure 10 (cont)

TTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCTCCGGCACCAAGTGGAACAAGAC
CCTGCAGCAGGTGGCCAAGAAGCTGCGCGAGCACTTCAACAACAAGACCATCATCTTCAAGCCCTCCTCCGGCGGCG
ACCTGGAGATCACCACCCACTCCTTCAACTGCCGCGGCGAGTTCTTCTACTGCAACACCTCCGGCCTGTTCAACTCC
ACCTGGATCGGCAACGGCACCAAGAACAACAACAACACCAACGACACCATCACCCTGCCCTGCCGCATCAAGCAGAT
CATCAACATGTGGCAGGGCGTGGGCCAGGCCATGTACGCCCCCCCCATCGAGGGCAAGATCACCTGCAAGTCCAACA
TCACCGGCCTGCTGCTGACCCGCGACGGCGGCAACAACAACACCAACGAGACCGAGATCTTCCGCCCCGGCGGCGGC
GACATGCGCGACAACTGGCGCTCCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCAC
CCGCTGCAAGCGCCGCGTGGTGGGCCGCCGCCGCCGCCGCCGCGCCGTGGGCATCGGCGCCGTGTTCCTGGGCTTCC
TGGGCGCCGCCGGCTCCACCATGGGCGCCGCCTCCATGACCCTGACCGTGCAGGCCCGCAACCTGCTGTCCGGCATC
GTGCAGCAGCAGTCCAACCTGCTGCGCGCCCCCGAGGCCCAGCAGCACCTGCTGAAGCTGACCGTGTGGGGCATCAA
GCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGCGCGACCAGCAGCTGCTGGGCATCTGGGGCTGCTCCG
GCAAGCTGATCTGCTGCACCAACGTGCCCTGGAACTCCTCCTGGTCCAACCGCAACCTGTCCGAGATCTGGGACAAC
ATGACCTGGCTGCAGTGGGACAAGGAGATCTCCAACTACACCCAGATCATCTACGGCCTGCTGGAGGAGTCCCAGAA
CCAGCAGGAGAAGAACGAGCAGGACCTGCTGGCCCTGGACGGCTCCGGCCTGAACGACATCTTCGAGGCCCAGAAGA
TCGAGTGGCACGAGTAGGGATCC

>HV1301260

GTCGACGCCACCATGGGCTCCCTGCAGCCCCTGGCCACCCTGTACCTGCTGGGCATGCTGGTGGCCTCCGTGCTGGC
CGCCGAGAACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCCAACACCACCCTGTTCTGCGCCT
CCGACGCCAAGGCCTACGACACCGAGGTGCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCC
CAGGAGATCGTGCTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAGGA
CATCATCTCCCTGTGGGACCAGTCCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTGACCCTGAACTGCACCA
ACGTGAACGTGACCAACACCACCAACAACACCGAGGAGAAGGGCGAGATCAAGAACTGCTCCTTCAACATCACCACC
GAGATCCGCGACAAGAAGCAGAAGGTGTACGCCCTGTTCTACCGCCTGGACGTGGTGCCCATCGACGACAACAACAA
CAACTCCTCCAACTACCGCCTGATCAACTGCAACACCTCCGCCATCACCCAGGCCTGCCCCAAGGTGTCCTTCGAGC
CCATCCCCATCCACTACTGCGCCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCACCGGC
CCCTGCAAGAACGTGTCCACCGTGCAGTGCACCCACGGCATCAAGCCCGTGGTGTCCACCCAGCTGCTGCTGAACGG
CTCCCTGGCCGAGGAGGAGATCATCATCCGCTCCGAGAACATCACCAACAACGCCAAGACCATCATCGTGCAGCTGA
ACGAGTCCGTGGAGATCAACTGCACCCGCCCCAACAACAACACCCGCAAGTCCATCCGCATCGGCCCCGGCCAGTGG
TTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCTCCGGCACCAAGTGGAACAAGAC
CCTGCAGCAGGTGGCCAAGAAGCTGCGCGAGCACTTCAACAACAAGACCATCATCTTCAAGCCCTCCTCCGGCGGCG
ACCTGGAGATCACCACCCACTCCTTCAACTGCCGCGGCGAGTTCTTCTACTGCAACACCTCCGGCCTGTTCAACTCC
ACCTGGATCGGCAACGGCACCAAGAACAACAACAACACCAACGACACCATCACCCTGCCCTGCCGCATCAAGCAGAT
CATCAACATGTGGCAGGGCGTGGGCCAGGCCATGTACGCCCCCCCCATCGAGGGCAAGATCACCTGCAAGTCCAACA
TCACCGGCCTGCTGCTGACCCGCGACGGCGGCAACAACAACACCAACGAGACCGAGATCTTCCGCCCCGGCGGCGGC
GACATGCGCGACAACTGGCGCTCCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCAC
CCGCTGCAAGCGCCGCGTGGTGGGCCGCCGCCGCCGCCGCCGCGCCGTGGGCATCGGCGCCGTGTTCCTGGGCTTCC
TGGGCGCCGCCGGCTCCACCATGGGCGCCGCCTCCATGACCCTGACCGTGCAGGCCCGCAACCTGCTGTCCGGCATC
GTGCAGCAGCAGTCCAACCTGCTGCGCGCCCCCGAGGCCCAGCAGCACCTGCTGAAGCTGACCGTGTGGGGCATCAA
GCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGCGCGACCAGCAGCTGCTGGGCATCTGGGGCTGCTCCG
GCAAGCTGATCTGCTGCACCAACGTGCCCTGGAACTCCTCCTGGTCCAACCGCAACCTGTCCGAGATCTGGGACAAC
ATGACCTGGCTGCAGTGGGACAAGGAGATCTCCAACTACACCCAGATCATCTACGGCCTGCTGGAGGAGTCCCAGAA
CCAGCAGGAGAAGAACGAGCAGGACCTGCTGGCCCTGGACGGCTCCGGCCTGAACGACATCTTCGAGGCCCAGAAGA
TCGAGTGGCACGAGTAGGGATCC

>HV1301521_ferritin

GTCACCGTCGTCGACGCTAGCACCATGGGCTCGCTCCAGCCGCTCGCGACGCTGTACCTCCTGGGCATGCTCGTGGC
GTCCGTGCTGGCGGCCGAGAACCTGTGGGTGACGGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCCAACACCACGC
TGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGTGCACAACGTGTGGGCGACCCACGCCTGCGTGCCGACG
GACCCCAACCCCCAGGAGATCGTGCTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCA
GATGCACGAGGACATCATCTCGCTGTGGGACCAGTCCCTGAAGCCGTGCGTGAAGCTGACGCCCCTGTGCGTGACCC
TGGACTGCACCAACGTGAAGGTGACGTCCACCACGTCCAACACGGAGGAGAAGGGGGAGATCAAGAACTGCTCCTTC
AACATCACCACCGAGATCCGCGACAAGAAGCAGAAGGTGTACGCGCTGTTCTACCGGCTGGACGTGGTGCCGATCGA
CGACAACAACAACAACTCCAGCAACTACCGCCTGATCAACTGCAACACCAGCGCCTGCACCCAGGCCTGCCCGAAGG
TGTCCTTCGAGCCCATCCCCATCCACTACTGCGCGCCGGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTTC

Figure 10 (cont)

AACGGCACCGGCCCCTGCAAGAACGTGTCCACCGTGCAGTGCACCCACGGGATCAAGCCCGTGGTGTCCACGCAGCT
GCTGCTGAACGGCTCCCTGGCCGAGGAGGAGATCATCATCCGCTCCGAGAACATCACGAACAACGCCAAGACCATCA
TCGTGCAGCTGAACGAGTCCGTGGAGATCAACTGCACCAGGCCCAACAACAACACCCGCAAGTCCATCCGGATCGGC
CCTGGCCAGGCGTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCGCACTGCAACATCTCGGGCACGAA
GTGGAACAAGACCCTGCAGCAGGTGGCGAAGAAGCTGCGCGAGCACTTCAACAACAAGACCATCATCTTCAAGCCCA
GCTCCGGCGGCGACCTGGAGATCACGACCCACTCCTTCAACTGCCGCGGCGAGTTCTTCTACTGCAACACCTCCGGC
CTGTTCAACTCGACGTGGATCGGGAACGGCACGAAGAACAACAACAACACCAACGACACCATCACCCTGCCCTGCCG
CATCAAGCAGATCATCAACATGTGGCAGGGCGTGGGCCAGTGCATGTACGCGCCGCCCATCGAGGGCAAGATCACCT
GCAAGTCCAACATCACCGGCCTGCTCCTGACGCGCGACGGCGGCAACAACAACACCAACGAGACCGAGATCTTCAGG
CCGGGCGGCGGCGACATGCGCGACAACTGGCGCTCGGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGG
CGTGGCGCCGACGCGCTGCAAGAGACGCGTGGTGGGCCGCAGACGAAGGAGACGGGCCGTGGGCATCGGCGCGGTGT
TCCTGGGCTTCCTGGGAGCAGCTGGTTCGACGATGGGCGCAGCTTCCATGACCCTGACAGTGCAGGCACGCAACCTG
CTCTCCGGCATCGTCCAGCAGCAGTCGAACCTGCTTCGAGCCCCCGAGGCGCAGCAGCACCTCCTCAAGCTGACCGT
GTGGGGCATCAAGCAGCTGCAGGCACGCGTGCTAGCCGTGGAGCGCTACCTCCGCGACCAGCAGCTGCTCGGAATCT
GGGGCTGCTCGGGCAAGCTGATCTGCTGCACCAACGTGCCGTGGAACAGCTCCTGGTCCAACCGCAACCTCTCGGAG
ATCTGGGACAACATGACCTGGCTCCAGTGGGACAAGGAGATCTCGAACTACACCCAGATCATCTACGGCCTGCTGGA
GGAGTCCCAGAACCAGCAGGAGAAGAACGAGCAGGACCTGCTGGCCCTGGACGGCGGAGGATCTGGCGACATTATCA
AGCTGCTGAACGAGCAAGTGAACAAAGAGATGAACAGCTCCAACCTGTACATGAGCATGAGCAGCTGGTGTTACACC
CACAGCCTTGATGGCGCCGGACTGTTCCTGTTTGATCACGCCGCCGAGGAATACGAGCACGCCAAGAAGCTGATCAT
CTTCCTGAACGAGAACAATGTGCCCGTGCAGCTGACCAGCATTAGCGCCCCAGAGCACAAGTTCGAGGGCCTGACAC
AGATCTTTCAGAAGGCCTACGAACACGAGCAGCACATCTCCGAGAGCATCAACAACATCGTGGACCACGCCATTAAG
AGCAAGGATCACGCCACCTTCAATTTTCTGCAGTGGTACGTGGCCGAACAGCACGAGGAAGAAGTGCTGTTCAAGGA
CATCCTGGACAAGATTGAGCTGATCGGCAACGAGAACCACGGCCTGTATCTGGCCGACCAGTACGTGAAGGGAATCG
CCAAGAGCCGGAAGTCCTGATAATCTAGAGGATCC

>**HV1301521**(originally HV1301405_N130D_N135K_N138S_N141S)

GTCGACGCTAGCACCATGGGCTCGCTCCAGCCGCTCGCGACGCTGTACCTCCTGGGCATGCTCGTGGCGTCCGTGCT
GGCGGCCGAGAACCTGTGGGTGACGGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCCAACACCACGCTGTTCTGCG
CCAGCGACGCCAAGGCCTACGACACCGAGGTGCACAACGTGTGGGCGACCCACGCCTGCGTGCCGACGGACCCCAAC
CCCCAGGAGATCGTGCTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGA
GGACATCATCTCGCTGTGGGACCAGTCCCTGAAGCCGTGCGTGAAGCTGACGCCCCTGTGCGTGACCCTGgACTGCA
CCAACGTGAAgGTGACGtcCACCACGtcCAACACGGAGGAGAAGGGGGAGATCAAGAACTGCTCCTTCAACATCACC
ACCGAGATCCGCGACAAGAAGCAGAAGGTGTACGCGCTGTTCTACCGGCTGGACGTGGTGCCGATCGACGACAACAA
CAACAACTCCAGCAACTACCGCCTGATCAACTGCAACACCAGCGCCTGCACCCAGGCCTGCCCGAAGGTGTCCTTCG
AGCCCATCCCCATCCACTACTGCGCGCCGGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCACC
GGCCCCTGCAAGAACGTGTCCACCGTGCAGTGCACCCACGGGATCAAGCCCGTGGTGTCCACGCAGCTGCTGCTGAA
CGGCTCCCTGGCCGAGGAGGAGATCATCATCCGCTCCGAGAACATCACGAACAACGCCAAGACCATCATCGTGCAGC
TGAACGAGTCCGTGGAGATCAACTGCACCAGGCCCAACAACAACACCCGCAAGTCCATCCGGATCGGCCCTGGCCAG
GCGTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCGCACTGCAACATCTCGGGCACGAAGTGGAACAA
GACCCTGCAGCAGGTGGCGAAGAAGCTGCGCGAGCACTTCAACAACAAGACCATCATCTTCAAGCCCAGCTCCGGCG
GCGACCTGGAGATCACGACCCACTCCTTCAACTGCCGCGGCGAGTTCTTCTACTGCAACACCTCCGGCCTGTTCAAC
TCGACGTGGATCGGGAACGGCACGAAGAACAACAACAACACCAACGACACCATCACCCTGCCCTGCCGCATCAAGCA
GATCATCAACATGTGGCAGGGCGTGGGCCAGTGCATGTACGCGCCGCCCATCGAGGGCAAGATCACCTGCAAGTCCA
ACATCACCGGCCTGCTCCTGACGCGCGACGGCGGCAACAACAACACCAACGAGACCGAGATCTTCAGGCCGGGCGGC
GGCGACATGCGCGACAACTGGCGCTCGGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCGCC
GACGCGCTGCAAGAGACGCGTGGTGGGCCGCAGACGAAGGAGACGGGCCGTGGGCATCGGCGCGGTGTTCCTGGGCT
TCCTGGGAGCAGCTGGTTCGACGATGGGCGCAGCTTCCATGACCCTGACAGTGCAGGCACGCAACCTGCTCTCCGGC
ATCGTCCAGCAGCAGTCGAACCTGCTTCGAGCCCCCGAGGCGCAGCAGCACCTCCTCAAGCTGACCGTGTGGGGCAT
CAAGCAGCTGCAGGCACGCGTGCTAGCCGTGGAGCGCTACCTCCGCGACCAGCAGCTGCTCGGAATCTGGGGCTGCT
CGGGCAAGCTGATCTGCTGCACCAACGTGCCGTGGAACAGCTCCTGGTCCAACCGCAACCTCTCGGAGATCTGGGAC
AACATGACCTGGCTCCAGTGGGACAAGGAGATCTCGAACTACACCCAGATCATCTACGGCCTGCTGGAGGAGTCCCA
GAACCAGCAGGAGAAGAACGAGCAGGACCTGCTGGCCCTGGACTGATAATCTAGAGGATCC

>**HV1301405_N138S_N141S**

GTCGACGCTAGCACCATGGGCTCGCTCCAGCCGCTCGCGACGCTGTACCTCCTGGGCATGCTCGTGGCGTCCGTGCT

Figure 10 (cont)

GGCGGCCGAGAACCTGTGGGTGACGGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCCAACACCACGCTGTTCTGCG
CCAGCGACGCCAAGGCCTACGACACCGAGGTGCACAACGTGTGGGCGACCCACGCCTGCGTGCCGACGGACCCCAAC
CCCCAGGAGATCGTGCTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGA
GGACATCATCTCGCTGTGGGACCAGTCCCTGAAGCCGTGCGTGAAGCTGACGCCCCTGTGCGTGACCCTGAACTGCA
CCAACGTGAACGTGACGtcCACCACGtcCAACACGGAGGAGAAGGGGGAGATCAAGAACTGCTCCTTCAACATCACC
ACCGAGATCCGCGACAAGAAGCAGAAGGTGTACGCGCTGTTCTACCGGCTGGACGTGGTGCCGATCGACGACAACAA
CAACAACTCCAGCAACTACCGCCTGATCAACTGCAACACCAGCGCCTGCACCCAGGCCTGCCCGAAGGTGTCCTTCG
AGCCCATCCCCATCCACTACTGCGCGCCGGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCACC
GGCCCCTGCAAGAACGTGTCCACCGTGCAGTGCACCCACGGGATCAAGCCCGTGGTGTCCACGCAGCTGCTGCTGAA
CGGCTCCCTGGCCGAGGAGGAGATCATCATCCGCTCCGAGAACATCACGAACAACGCCAAGACCATCATCGTGCAGC
TGAACGAGTCCGTGGAGATCAACTGCACCAGGCCCAACAACAACACCCGCAAGTCCATCCGGATCGGCCCTGGCCAG
GCGTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCGCACTGCAACATCTCGGGCACGAAGTGGAACAA
GACCCTGCAGCAGGTGGCGAAGAAGCTGCGCGAGCACTTCAACAACAAGACCATCATCTTCAAGCCCAGCTCCGGCG
GCGACCTGGAGATCACGACCCACTCCTTCAACTGCCGCGGCGAGTTCTTCTACTGCAACACCTCCGGCCTGTTCAAC
TCGACGTGGATCGGGAACGGCACGAAGAACAACAACAACACCAACGACACCATCACCCTGCCCTGCCGCATCAAGCA
GATCATCAACATGTGGCAGGGCGTGGGCCAGTGCATGTACGCGCCGCCCATCGAGGGCAAGATCACCTGCAAGTCCA
ACATCACCGGCCTGCTCCTGACGCGCGACGGCGGCAACAACAACACCAACGAGACCGAGATCTTCAGGCCGGGCGGC
GGCGACATGCGCGACAACTGGCGCTCGGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCGCC
GACGCGCTGCAAGAGACGCGTGGTGGGCCGCAGACGAAGGAGACGGGCCGTGGGCATCGGCGCGGTGTTCCTGGGCT
TCCTGGGAGCAGCTGGTTCGACGATGGGCGCAGCTTCCATGACCCTGACAGTGCAGGCACGCAACCTGCTCTCCGGC
ATCGTCCAGCAGCAGTCGAACCTGCTTCGAGCCCCCGAGGCGCAGCAGCACCTCCTCAAGCTGACCGTGTGGGGCAT
CAAGCAGCTGCAGGCACGCGTGCTAGCCGTGGAGCGCTACCTCCGCGACCAGCAGCTGCTCGGAATCTGGGGCTGCT
CGGGCAAGCTGATCTGCTGCACCAACGTGCCGTGGAACAGCTCCTGGTCCAACCGCAACCTCTCGGAGATCTGGGAC
AACATGACCTGGCTCCAGTGGGACAAGGAGATCTCGAACTACACCCAGATCATCTACGGCCTGCTGGAGGAGTCCCA
GAACCAGCAGGAGAAGAACGAGCAGGACCTGCTGGCCCTGGACTGATAATCTAGAGGATCC

>HV1301405

GTCGACGCTAGCACCATGGGCTCGCTCCAGCCGCTCGCGACGCTGTACCTCCTGGGCATGCTCGTGGCGTCCGTGCT
GGCGGCCGAGAACCTGTGGGTGACGGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCCAACACCACGCTGTTCTGCG
CCAGCGACGCCAAGGCCTACGACACCGAGGTGCACAACGTGTGGGCGACCCACGCCTGCGTGCCGACGGACCCCAAC
CCCCAGGAGATCGTGCTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGA
GGACATCATCTCGCTGTGGGACCAGTCCCTGAAGCCGTGCGTGAAGCTGACGCCCCTGTGCGTGACCCTGAACTGCA
CCAACGTGAACGTGACGAACACCACGAACAACACGGAGGAGAAGGGGGAGATCAAGAACTGCTCCTTCAACATCACC
ACCGAGATCCGCGACAAGAAGCAGAAGGTGTACGCGCTGTTCTACCGGCTGGACGTGGTGCCGATCGACGACAACAA
CAACAACTCCAGCAACTACCGCCTGATCAACTGCAACACCAGCGCCTGCACCCAGGCCTGCCCGAAGGTGTCCTTCG
AGCCCATCCCCATCCACTACTGCGCGCCGGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCACC
GGCCCCTGCAAGAACGTGTCCACCGTGCAGTGCACCCACGGGATCAAGCCCGTGGTGTCCACGCAGCTGCTGCTGAA
CGGCTCCCTGGCCGAGGAGGAGATCATCATCCGCTCCGAGAACATCACGAACAACGCCAAGACCATCATCGTGCAGC
TGAACGAGTCCGTGGAGATCAACTGCACCAGGCCCAACAACAACACCCGCAAGTCCATCCGGATCGGCCCTGGCCAG
GCGTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCGCACTGCAACATCTCGGGCACGAAGTGGAACAA
GACCCTGCAGCAGGTGGCGAAGAAGCTGCGCGAGCACTTCAACAACAAGACCATCATCTTCAAGCCCAGCTCCGGCG
GCGACCTGGAGATCACGACCCACTCCTTCAACTGCCGCGGCGAGTTCTTCTACTGCAACACCTCCGGCCTGTTCAAC
TCGACGTGGATCGGGAACGGCACGAAGAACAACAACAACACCAACGACACCATCACCCTGCCCTGCCGCATCAAGCA
GATCATCAACATGTGGCAGGGCGTGGGCCAGTGCATGTACGCGCCGCCCATCGAGGGCAAGATCACCTGCAAGTCCA
ACATCACCGGCCTGCTCCTGACGCGCGACGGCGGCAACAACAACACCAACGAGACCGAGATCTTCAGGCCGGGCGGC
GGCGACATGCGCGACAACTGGCGCTCGGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCGCC
GACGCGCTGCAAGAGACGCGTGGTGGGCCGCAGACGAAGGAGACGGGCCGTGGGCATCGGCGCGGTGTTCCTGGGCT
TCCTGGGAGCAGCTGGTTCGACGATGGGCGCAGCTTCCATGACCCTGACAGTGCAGGCACGCAACCTGCTCTCCGGC
ATCGTCCAGCAGCAGTCGAACCTGCTTCGAGCCCCCGAGGCGCAGCAGCACCTCCTCAAGCTGACCGTGTGGGGCAT
CAAGCAGCTGCAGGCACGCGTGCTAGCCGTGGAGCGCTACCTCCGCGACCAGCAGCTGCTCGGAATCTGGGGCTGCT
CGGGCAAGCTGATCTGCTGCACCAACGTGCCGTGGAACAGCTCCTGGTCCAACCGCAACCTCTCGGAGATCTGGGAC
AACATGACCTGGCTCCAGTGGGACAAGGAGATCTCGAACTACACCCAGATCATCTACGGCCTGCTGGAGGAGTCCCA
GAACCAGCAGGAGAAGAACGAGCAGGACCTGCTGGCCCTGGACTGATAATCTAGAGGATCC

Figure 10 (cont)

>HV1301258_N301A

gtcgacgccacc**ATG**GGCTCCCTGCAGCCCCTGGCCACCCTGTACCTGCTGGGCATGCTGGTGGCCTCCGTGCTGGC
CGCCGAGAACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCCAACACCACCCTGTTCTGCGCCT
CCGACGCCAAGGCCTACGACACCGAGGTGCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCC
CAGGAGATCGTGCTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAGGA
CATCATCTCCCTGTGGGACCAGTCCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTGACCCTGAACTGCACCA
ACGTGAACGTGACCAACACCACCAACAACACCGAGGAGAAGGGCGAGATCAAGAACTGCTCCTTCAACATCACCACC
GAGATCCGCGACAAGAAGCAGAAGGTGTACGCCCTGTTCTACCGCCTGGACGTGGTGCCCATCGACGACAACAACAA
CAACTCCTCCAACTACCGCCTGATCAACTGCAACACCTCCGCCTGCACCCAGGCCTGCCCCAAGGTGTCCTTCGAGC
CCATCCCCATCCACTACTGCGCCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCACCGGC
CCCTGCAAGAACGTGTCCACCGTGCAGTGCACCCACGGCATCAAGCCCGTGGTGTCCACCCAGCTGCTGCTGAACGG
CTCCCTGGCCGAGGAGGAGATCATCATCCGCTCCGAGAACATCACCAACAACGCCAAGACCATCATCGTGCAGCTGA
ACGAGTCCGTGGAGATCAACTGCACCCGCCCCAACGCCAACACCCGCAAGTCCATCCGCATCGGCCCCGGCCAGGCC
TTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCTCCGGCACCAAGTGGAACAAGAC
CCTGCAGCAGGTGGCCAAGAAGCTGCGCGAGCACTTCAACAACAAGACCATCATCTTCAAGCCCTCCTCCGGCGGCG
ACCTGGAGATCACCACCCACTCCTTCAACTGCCGCGGCGAGTTCTTCTACTGCAACACCTCCGGCCTGTTCAACTCC
ACCTGGATCGGCAACGGCACCAAGAACAACAACAACACCAACGACACCATCACCCTGCCCTGCCGCATCAAGCAGAT
CATCAACATGTGGCAGGGCGTGGGCCAGTGCATGTACGCCCCCCCCATCGAGGGCAAGATCACCTGCAAGTCCAACA
TCACCGGCCTGCTGCTGACCCGCGACGGCGGCAACAACAACACCAACGAGACCGAGATCTTCCGCCCCGGCGGCGGC
GACATGCGCGACAACTGGCGCTCCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCAC
CCGCTGCAAGCGCCGCGTGGTGGGCCGCCGCCGCCGCCGCCGCCGCCGTGGGCATCGGCGCCGTGTTCCTGGGCTTCC
TGGGCGCCGCCGGCTCCACCATGGGCGCCGCCTCCATGACCCTGACCGTGCAGGCCCGCAACCTGCTGTCCGGCATC
GTGCAGCAGCAGTCCAACCTGCTGCGCGCCCCCGAGGCCCAGCAGCACCTGCTGAAGCTGACCGTGTGGGGCATCAA
GCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGCGCGACCAGCAGCTGCTGGGCATCTGGGGCTGCTCCG
GCAAGCTGATCTGCTGCACCAACGTGCCCTGGAACTCCTCCTGGTCCAACCGCAACCTGTCCGAGATCTGGGACAAC
ATGACCTGGCTGCAGTGGGACAAGGAGATCTCCAACTACACCCAGATCATCTACGGCCTGCTGGAGGAGTCCCAGAA
CCAGCAGGAGAAGAACGAGCAGGACCTGCTGGCCCTGGACGGCTCCGGCCTGAACGACATCTTCGAGGCCCAGAAGA
TCGAGTGGCACGAG**TAG**ggatcc

>HV1301258_N332A, CON-Schim.6R.DS.SOSIP.664_avi_N332A

gtcgacgccacc**ATG**GGCTCCCTGCAGCCCCTGGCCACCCTGTACCTGCTGGGCATGCTGGTGGCCTCCGTGCTGGC
CGCCGAGAACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCCAACACCACCCTGTTCTGCGCCT
CCGACGCCAAGGCCTACGACACCGAGGTGCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCC
CAGGAGATCGTGCTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAGGA
CATCATCTCCCTGTGGGACCAGTCCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTGACCCTGAACTGCACCA
ACGTGAACGTGACCAACACCACCAACAACACCGAGGAGAAGGGCGAGATCAAGAACTGCTCCTTCAACATCACCACC
GAGATCCGCGACAAGAAGCAGAAGGTGTACGCCCTGTTCTACCGCCTGGACGTGGTGCCCATCGACGACAACAACAA
CAACTCCTCCAACTACCGCCTGATCAACTGCAACACCTCCGCCTGCACCCAGGCCTGCCCCAAGGTGTCCTTCGAGC
CCATCCCCATCCACTACTGCGCCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCACCGGC
CCCTGCAAGAACGTGTCCACCGTGCAGTGCACCCACGGCATCAAGCCCGTGGTGTCCACCCAGCTGCTGCTGAACGG
CTCCCTGGCCGAGGAGGAGATCATCATCCGCTCCGAGAACATCACCAACAACGCCAAGACCATCATCGTGCAGCTGA
ACGAGTCCGTGGAGATCAACTGCACCCGCCCCAACAACAACACCCGCAAGTCCATCCGCATCGGCCCCGGCCAGGCC
TTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCTCCGGCACCAAGTGGAACAAGAC
CCTGCAGCAGGTGGCCAAGAAGCTGCGCGAGCACTTCAACAACAAGACCATCATCTTCAAGCCCTCCTCCGGCGGCG
ACCTGGAGATCACCACCCACTCCTTCAACTGCCGCGGCGAGTTCTTCTACTGCAACACCTCCGGCCTGTTCAACTCC
ACCTGGATCGGCAACGGCACCAAGAACAACAACAACACCAACGACACCATCACCCTGCCCTGCCGCATCAAGCAGAT
CATCAACATGTGGCAGGGCGTGGGCCAGTGCATGTACGCCCCCCCCATCGAGGGCAAGATCACCTGCAAGTCCAACA
TCACCGGCCTGCTGCTGACCCGCGACGGCGGCAACAACAACACCAACGAGACCGAGATCTTCCGCCCCGGCGGCGGC
GACATGCGCGACAACTGGCGCTCCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCAC
CCGCTGCAAGCGCCGCGTGGTGGGCCGCCGCCGCCGCCGCCGCCGTGGGCATCGGCGCCGTGTTCCTGGGCTTCC
TGGGCGCCGCCGGCTCCACCATGGGCGCCGCCTCCATGACCCTGACCGTGCAGGCCCGCAACCTGCTGTCCGGCATC
GTGCAGCAGCAGTCCAACCTGCTGCGCGCCCCCGAGGCCCAGCAGCACCTGCTGAAGCTGACCGTGTGGGGCATCAA
GCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGCGCGACCAGCAGCTGCTGGGCATCTGGGGCTGCTCCG
GCAAGCTGATCTGCTGCACCAACGTGCCCTGGAACTCCTCCTGGTCCAACCGCAACCTGTCCGAGATCTGGGACAAC

Figure 10 (cont)

ATGACCTGGCTGCAGTGGGACAAGGAGATCTCCAACTACACCCAGATCATCTACGGCCTGCTGGAGGAGTCCCAGAA
CCAGCAGGAGAAGAACGAGCAGGACCTGCTGGCCCTGGACGGCTCCGGCCTGAACGACATCTTCGAGGCCCAGAAGA
TCGAGTGGCACGAG**tag**ggatcc
>HV1301521_c_sorta
GTCGACGCTAGCACCATGGGCTCGCTCCAGCCGCTCGCGACGCTGTACCTCCTGGGCATGCTCGTGGCGTCCGTGCT
GGCGGCCGAGAACCTGTGGGTGACGGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCCAACACCACGCTGTTCTGCG
CCAGCGACGCCAAGGCCTACGACACCGAGGTGCACAACGTGTGGGCGACCCACGCCTGCGTGCCGACGGACCCCAAC
CCCCAGGAGATCGTGCTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGA
GGACATCATCTCGCTGTGGGACCAGTCCCTGAAGCCGTGCGTGAAGCTGACGCCCCTGTGCGTGACCCTGGACTGCA
CCAACGTGAAGGTGACGTCCACCACGTCCAACACGGAGGAGAAGGGGGAGATCAAGAACTGCTCCTTCAACATCACC
ACCGAGATCCGCGACAAGAAGCAGAAGGTGTACGCGCTGTTCTACCGGCTGGACGTGGTGCCGATCGACGACAACAA
CAACAACTCCAGCAACTACCGCCTGATCAACTGCAACACCAGCGCCTGCACCCAGGCCTGCCCGAAGGTGTCCTTCG
AGCCCATCCCCATCCACTACTGCGCGCCGGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCACC
GGCCCCTGCAAGAACGTGTCCACCGTGCAGTGCACCCACGGGATCAAGCCCGTGGTGTCCACGCAGCTGCTGCTGAA
CGGCTCCCTGGCCGAGGAGGAGATCATCATCCGCTCCGAGAACATCACGAACAACGCCAAGACCATCATCGTGCAGC
TGAACGAGTCCGTGGAGATCAACTGCACCAGGCCCAACAACAACACCCGCAAGTCCATCCGGATCGGCCCTGGCCAG
GCGTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCGCACTGCAACATCTCGGGCACGAAGTGGAACAA
GACCCTGCAGCAGGTGGCGAAGAAGCTGCGCGAGCACTTCAACAACAAGACCATCATCTTCAAGCCCAGCTCCGGCG
GCGACCTGGAGATCACGACCCACTCCTTCAACTGCCGCGGCGAGTTCTTCTACTGCAACACCTCCGGCCTGTTCAAC
TCGACGTGGATCGGGAACGGCACGAAGAACAACAACAACACCAACGACACCATCACCCTGCCCTGCCGCATCAAGCA
GATCATCAACATGTGGCAGGGCGTGGGCCAGTGCATGTACGCGCCGCCCATCGAGGGCAAGATCACCTGCAAGTCCA
ACATCACCGGCCTGCTCCTGACGCGCGACGGCGGCAACAACAACACCAACGAGACCGAGATCTTCAGGCCGGGCGGC
GGCGACATGCGCGACAACTGGCGCTCGGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCGCC
GACGCGCTGCAAGAGACGCGTGGTGGGCCGCAGACGAAGGAGACGGGCCGTGGGCATCGGCGCGGTGTTCCTGGGCT
TCCTGGGAGCAGCTGGTTCGACGATGGGCGCAGCTTCCATGACCCTGACAGTGCAGGCACGCAACCTGCTCTCCGGC
ATCGTCCAGCAGCAGTCGAACCTGCTTCGAGCCCCCGAGGCGCAGCAGCACCTCCTCAAGCTGACCGTGTGGGGCAT
CAAGCAGCTGCAGGCACGCGTGCTAGCCGTGGAGCGCTACCTCCGCGACCAGCAGCTGCTCGGAATCTGGGGCTGCT
CGGGCAAGCTGATCTGCTGCACCAACGTGCCGTGGAACAGCTCCTGGTCCAACCGCAACCTCTCGGAGATCTGGGAC
AACATGACCTGGCTCCAGTGGGACAAGGAGATCTCGAACTACACCCAGATCATCTACGGCCTGCTGGAGGAGTCCCA
GAACCAGCAGGAGAAGAACGAGCAGGACCTGCTGGCCCTGGACCTGCCTAGCACCGGAGGATGATAATCTAGAGGAT
CC
>HV1301639_avi
ATGGGCTCCCTGCAGCCCCTGGCCACCCTGTACCTGCTGGGCATGCTGGTGGCCTCCGTGCTGGCCGCCGAGAACCT
GTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCCAACACCACCCTGTTCTGCGCCTCCGACGCCAAGG
CCTACGACACCGAGGTGCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTG
CTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAGGACATCATCTCCCT
GTGGGACCAGTCCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTGACCCTGGACTGCACCAACGTGAAGGTGA
CCAACACCACCAACAACACCGAGGAGAAGGGCGAGATCAAGAACTGCTCCTTCAACATCACCACCGAGATCCGCGAC
AAGAAGCAGAAGGTGTACGCCCTGTTCTACCGCCTGGACGTGGTGCCCATCGACGACAACAACAACAACTCCTCCAA
CTACCGCCTGATCAACTGCAACACCTCCGCCTGCACCCAGGCCTGCCCCAAGGTGTCCTTCGAGCCCATCCCCATCC
ACTACTGCGCCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCACCGGCCCCTGCAAGAAC
GTGTCCACCGTGCAGTGCACCCACGGCATCAAGCCCGTGGTGTCCACCCAGCTGCTGCTGAACGGCTCCCTGGCCGA
GGAGGAGATCATCATCCGCTCCGAGAACATCACCAACAACGCCAAGACCATCATCGTGCAGCTGAACGAGTCCGTGG
AGATCAACTGCACCCGCCCCAACAACAACACCCGCAAGTCCATCCGCATCGGCCCCGGCCAGGCCTTCTACGCCACC
GGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCTCCGGCACCAAGTGGAACAAGACCCTGCAGCAGGT
GGCCAAGAAGCTGCGCGAGCACTTCAACAACAAGACCATCATCTTCAAGCCCTCCTCCGGCGGCGACCTGGAGATCA
CCACCCACTCCTTCAACTGCCGCGGCGAGTTCTTCTACTGCAACACCTCCGGCCTGTTCAACTCCACCTGGATCGGC
AACGGCACCAAGAACAACAACAACACCAACGACACCATCACCCTGCCCTGCCGCATCAAGCAGATCATCAACATGTG
GCAGGGCGTGGGCCAGTGCATGTACGCCCCCCCCATCGAGGGCAAGATCACCTGCAAGTCCAACATCACCGGCCTGC
TGCTGACCCGCGACGGCGGCAACAACAACACCAACGAGACCGAGATCTTCCGCCCCGGCGGCGGCGACATGCGCGAC
AACTGGCGCTCCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCCGCTGCAAGCG
CCGCGTGGTGGGCCGCCGCCGCCGCCGCCGCGCCGTGGGCATCGGCGCCGTGTTCCTGGGCTTCCTGGGCGCCGCCG
GCTCCACCATGGGCGCCGCCTCCATGACCCTGACCGTGCAGGCCCGCAACCTGCTGTCCGGCATCGTGCAGCAGCAG
TCCAACCTGCTGCGCGCCCCCGAGGCCCAGCAGCACCTGCTGAAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGC

Figure 10 (cont)

CCGCGTGCTGGCCGTGGAGCGCTACCTGCGCGACCAGCAGCTGCTGGGCATCTGGGGCTGCTCCGGCAAGCTGATCT
GCTGCACCAACGTGCCCTGGAACTCCTCCTGGTCCAACCGCAACCTGTCCGAGATCTGGGACAACATGACCTGGCTG
CAGTGGGACAAGGAGATCTCCAACTACACCCAGATCATCTACGGCCTGCTGGAGGAGTCCCAGAACCAGCAGGAGAA
GAACGAGCAGGACCTGCTGGCCCTGGACGGCTCCGGCCTGAACGACATCTTCGAGGCCCAGAAGATCGAGTGGCACG
AGTAGGGATCC

>HV1301640_avi
ATGGGCTCCCTGCAGCCCCTGGCCACCCTGTACCTGCTGGGCATGCTGGTGGCCTCCGTGCTGGCCGCCGAGAACCT
GTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCCAACACCACCCTGTTCTGCGCCTCCGACGCCAAGG
CCTACGACACCGAGGTGCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTG
CTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAGGACATCATCTCCCT
GTGGGACCAGTCCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTGACCCTGAACTGCACCAACGTGAACGTGA
CCAGCACCACCAGCAACACCGAGGAGAAGGGCGAGATCAAGAACTGCTCCTTCAACATCACCACCGAGATCCGCGAC
AAGAAGCAGAAGGTGTACGCCCTGTTCTACCGCCTGGACGTGGTGCCCATCGACGACAACAACAACAACTCCTCCAA
CTACCGCCTGATCAACTGCAACACCTCCGCCTGCACCCAGGCCTGCCCCAAGGTGTCCTTCGAGCCCATCCCCATCC
ACTACTGCGCCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCACCGGCCCCTGCAAGAAC
GTGTCCACCGTGCAGTGCACCCACGGCATCAAGCCCGTGGTGTCCACCCAGCTGCTGCTGAACGGCTCCCTGGCCGA
GGAGGAGATCATCATCCGCTCCGAGAACATCACCAACAACGCCAAGACCATCATCGTGCAGCTGAACGAGTCCGTGG
AGATCAACTGCACCCGCCCCAACAACAACACCCGCAAGTCCATCCGCATCGGCCCCGGCCAGGCCTTCTACGCCACC
GGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCTCCGGCACCAAGTGGAACAAGACCCTGCAGCAGGT
GGCCAAGAAGCTGCGCGAGCACTTCAACAACAAGACCATCATCTTCAAGCCCTCCTCCGGCGGCGACCTGGAGATCA
CCACCCACTCCTTCAACTGCCGCGGCGAGTTCTTCTACTGCAACACCTCCGGCCTGTTCAACTCCACCTGGATCGGC
AACGGCACCAAGAACAACAACAACACCAACGACACCATCACCCTGCCCTGCCGCATCAAGCAGATCATCAACATGTG
GCAGGGCGTGGGCCAGTGCATGTACGCCCCCCCCATCGAGGGCAAGATCACCTGCAAGTCCAACATCACCGGCCTGC
TGCTGACCCGCGACGGCGGCAACAACAACACCAACGAGACCGAGATCTTCCGCCCCGGCGGCGGCGACATGCGCGAC
AACTGGCGCTCCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCCGCTGCAAGCG
CCGCGTGGTGGGCCGCCGCCGCCGCCGCCGCGCCGTGGGCATCGGCGCCGTGTTCCTGGGCTTCCTGGGCGCCGCCG
GCTCCACCATGGGCGCCGCCTCCATGACCCTGACCGTGCAGGCCCGCAACCTGCTGTCCGGCATCGTGCAGCAGCAG
TCCAACCTGCTGCGCGCCCCCGAGGCCCAGCAGCACCTGCTGAAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGC
CCGCGTGCTGGCCGTGGAGCGCTACCTGCGCGACCAGCAGCTGCTGGGCATCTGGGGCTGCTCCGGCAAGCTGATCT
GCTGCACCAACGTGCCCTGGAACTCCTCCTGGTCCAACCGCAACCTGTCCGAGATCTGGGACAACATGACCTGGCTG
CAGTGGGACAAGGAGATCTCCAACTACACCCAGATCATCTACGGCCTGCTGGAGGAGTCCCAGAACCAGCAGGAGAA
GAACGAGCAGGACCTGCTGGCCCTGGACGGCTCCGGCCTGAACGACATCTTCGAGGCCCAGAAGATCGAGTGGCACG
AGTAGGGATCC

>HV1301641_avi
ATGGGCTCCCTGCAGCCCCTGGCCACCCTGTACCTGCTGGGCATGCTGGTGGCCTCCGTGCTGGCCGCCGAGAACCT
GTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCCAACACCACCCTGTTCTGCGCCTCCGACGCCAAGG
CCTACGACACCGAGGTGCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTG
CTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAGGACATCATCTCCCT
GTGGGACCAGTCCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTGACCCTGGACTGCACCAACGTGAAGGTGA
CCAGCACCACCAGCAACACCGAGGAGAAGGGCGAGATCAAGAACTGCTCCTTCAACATCACCACCGAGATCCGCGAC
AAGAAGCAGAAGGTGTACGCCCTGTTCTACCGCCTGGACGTGGTGCCCATCGACGACAACAACAACAACTCCTCCAA
CTACCGCCTGATCAACTGCAACACCTCCGCCTGCACCCAGGCCTGCCCCAAGGTGTCCTTCGAGCCCATCCCCATCC
ACTACTGCGCCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCACCGGCCCCTGCAAGAAC
GTGTCCACCGTGCAGTGCACCCACGGCATCAAGCCCGTGGTGTCCACCCAGCTGCTGCTGAACGGCTCCCTGGCCGA
GGAGGAGATCATCATCCGCTCCGAGAACATCACCAACAACGCCAAGACCATCATCGTGCAGCTGAACGAGTCCGTGG
AGATCAACTGCACCCGCCCCAACAACAACACCCGCAAGTCCATCCGCATCGGCCCCGGCCAGGCCTTCTACGCCACC
GGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCTCCGGCACCAAGTGGAACAAGACCCTGCAGCAGGT
GGCCAAGAAGCTGCGCGAGCACTTCAACAACAAGACCATCATCTTCAAGCCCTCCTCCGGCGGCGACCTGGAGATCA
CCACCCACTCCTTCAACTGCCGCGGCGAGTTCTTCTACTGCAACACCTCCGGCCTGTTCAACTCCACCTGGATCGGC
AACGGCACCAAGAACAACAACAACACCAACGACACCATCACCCTGCCCTGCCGCATCAAGCAGATCATCAACATGTG
GCAGGGCGTGGGCCAGTGCATGTACGCCCCCCCCATCGAGGGCAAGATCACCTGCAAGTCCAACATCACCGGCCTGC
TGCTGACCCGCGACGGCGGCAACAACAACACCAACGAGACCGAGATCTTCCGCCCCGGCGGCGGCGACATGCGCGAC

Figure 10 (cont)

AACTGGCGCTCCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCCGCTGCAAGCG
CCGCGTGGTGGGCCGCCGCCGCCGCCGCCGCGCCGTGGGCATCGGCGCCGTGTTCCTGGGCTTCCTGGGCGCCGCCG
GCTCCACCATGGGCGCCGCCTCCATGACCCTGACCGTGCAGGCCCGCAACCTGCTGTCCGGCATCGTGCAGCAGCAG
TCCAACCTGCTGCGCGCCCCCGAGGCCCAGCAGCACCTGCTGAAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGC
CCGCGTGCTGGCCGTGGAGCGCTACCTGCGCGACCAGCAGCTGCTGGGCATCTGGGGCTGCTCCGGCAAGCTGATCT
GCTGCACCAACGTGCCCTGGAACTCCTCCTGGTCCAACCGCAACCTGTCCGAGATCTGGGACAACATGACCTGGCTG
CAGTGGGACAAGGAGATCTCCAACTACACCCAGATCATCTACGGCCTGCTGGAGGAGTCCCAGAACCAGCAGGAGAA
GAACGAGCAGGACCTGCTGGCCCTGGACGGCTCCGGCCTGAACGACATCTTCGAGGCCCAGAAGATCGAGTGGCACG
AGTAGGGATCC

>HV1301613
ATGGGCTCGCTCCAGCCGCTCGCGACGCTGTACCTCCTGGGCATGCTCGTGGCGTCCGTGCTGGCGGCCGAGAACCT
GTGGGTGACGGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCCAACACCACGCTGTTCTGCGCCAGCGACGCCAAGG
CCTACGACACCAAGGTGCACAACGTGTGGGCGACCCACGCCTGCGTGCCGACGGACCCCAACCCCCAGGAGATCGTG
CTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAGGACATCATCTCGCT
GTGGGACCAGTCCCTGAAGCCGTGCGTGAAGCTGACGCCCCTGTGCGTGACCCTGAACTGCACCAACGTGAACGTGA
CGAACACCACGAACAACACGGAGGAGAAGGGGGAGATCAAGAACTGCTCCTTCAACATCACCACCGAGATCCGCGAC
AAGAAGCAGAAGGTGTACGCGCTGTTCTACCGGCTGGACGTGGTGCCGATCGACGACAACAACAACAACTCCAGCAA
CTACCGCCTGATCAACTGCAACACCAGCGCCTGCACCCAGGCCTGCCCGAAGGTGTCCTTCGAGCCCATCCCCATCC
ACTACTGCGCGCCGGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCACCGGCCCCTGCAAGAAC
GTGTCCACCGTGCAGTGCACCCACGGGATCAAGCCCGTGGTGTCCACGCAGCTGCTGCTGAACGGCTCCCTGGCCGA
GGAGGAGATCATCATCCGCTCCGAGAACATCACGAACAACGCCAAGACCATCATCGTGCAGCTGAACGAGTCCGTGG
AGATCAACTGCACCAGGCCCAACAACAACACCCGCAAGTCCATCCGGATCGGCCCTGGCCAGTGGTTCTACGCCACC
GGCGACATCATCGGCGACATCCGCCAGGCGCACTGCAACATCTCGGGCACGAAGTGGAACAAGACCCTGCAGCAGGT
GGCGAAGAAGCTGCGCGAGCACTTCAACAACAAGACCATCATCTTCAAGCCCAGCTCCGGCGGCGACCTGGAGATCA
CGACCCACTCCTTCAACTGCCGCGGCGAGTTCTTCTACTGCAACACCTCCGGCCTGTTCAACTCGACGTGGATCGGG
AACGGCACGAAGAACAACAACAACACCAACGACACCATCACCCTGCCCTGCCGCATCAAGCAGATCATCAACATGTG
GCAGGGCGTGGGCCAGTGCATGTACGCGCCGCCCATCGAGGGCAAGATCACCTGCAAGTCCAACATCACCGGCCTGC
TCCTGACGCGCGACGGCGGCAACAACAACACCAACGAGACCGAGATCTTCAGGCCGGGCGGCGGCGACATGCGCGAC
AACTGGCGCTCGGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCGCCGACGCGCTGCAAGAG
ACGCGTGGTGGGCCGCAGACGAAGGAGACGGGCCGTGGGCATCGGCGCGGTGTTCCTGGGCTTCCTGGGAGCAGCTG
GTTCGACGATGGGCGCAGCTTCCATGACCCTGACAGTGCAGGCACGCAACCTGCTCTCCGGCATCGTCCAGCAGCAG
TCGAACCTGCTTCGAGCCCCCGAGGCGCAGCAGCACCTCCTCAAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGC
ACGCGTGCTAGCCGTGGAGCGCTACCTCCGCGACCAGCAGCTGCTCGGAATCTGGGGCTGCTCGGGCAAGCTGATCT
GCTGCACCAACGTGCCGTGGAACAGCTCCTGGTCCAACCGCAACCTCTCGGAGATCTGGGACAACATGACCTGGCTC
CAGTGGGACAAGGAGATCTCGAACTACACCCAGATCATCTACGGCCTGCTGGAGGAGTCCCAGAACCAGCAGGAGAA
GAACGAGCAGGACCTGCTGGCCCTGGACTGATAATCTAGAGGATCCAG

>HV130111_N137A CON-Sgp140CFI_avi N137A
AAGCTTGTCGACACCATGCGCGTGCGCGGCATCCAGCGCAACTGCCAGCACCTGTGGCGCTGGGGCACCCTGATCCT
GGGCATGCTGATGATCTGCTCCGCCGCCGAGAACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGG
CCAACACCACCCTGTTCTGCGCCTCCGACGCCAAGGCCTACGACACCGAGGTGCACAACGTGTGGGCCACCCACGCC
TGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAA
CATGGTGGAGCAGATGCACGAGGACATCATCTCCCTGTGGGACCAGTCCCTGAAGCCCTGCGTGAAGCTGACCCCCC
TGTGCGTGACCCTGAACTGCACCAACGTGAACGTGACCGCCACCACCAACAACACCGAGGAGAAGGGCGAGATCAAG
AACTGCTCCTTCAACATCACCACCGAGATCCGCGACAAGAAGCAGAAGGTGTACGCCCTGTTCTACCGCCTGGACGT
GGTGCCCATCGACGACAACAACAACAACTCCTCCAACTACCGCCTGATCAACTGCAACACCTCCGCCATCACCCAGG
CCTGCCCCAAGGTGTCCTTCGAGCCCATCCCCATCCACTACTGCGCCCCCGCCGGCTTCGCCATCCTGAAGTGCAAC
GACAAGAAGTTCAACGGCACCGGCCCCTGCAAGAACGTGTCCACCGTGCAGTGCACCCACGGCATCAAGCCCGTGGT
GTCCACCCAGCTGCTGCTGAACGGCTCCCTGGCCGAGGAGGAGATCATCATCCGCTCCGAGAACATCACCAACAACG
CCAAGACCATCATCGTGCAGCTGAACGAGTCCGTGGAGATCAACTGCACCCGCCCCAACAACAACACCCGCAAGTCC
ATCCGCATCGGCCCCGGCCAGGCCTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACAT
CTCCGGCACCAAGTGGAACAAGACCCTGCAGCAGGTGGCCAAGAAGCTGCGCGAGCACTTCAACAACAAGACCATCA
TCTTCAAGCCCTCCTCCGGCGGCGACCTGGAGATCACCACCCACTCCTTCAACTGCCGCGGCGAGTTCTTCTACTGC

Figure 10 (cont)

AACACCTCCGGCCTGTTCAACTCCACCTGGATCGGCAACGGCACCAAGAACAACAACAACACCAACGACACCATCAC
CCTGCCCTGCCGCATCAAGCAGATCATCAACATGTGGCAGGGCGTGGGCCAGGCCATGTACGCCCCCCCCATCGAGG
GCAAGATCACCTGCAAGTCCAACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAACAACAACACCAACGAGACC
GAGATCTTCCGCCCCGGCGGCGGCGACATGCGCGACAACTGGCGCTCCGAGCTGTACAAGTACAAGGTGGTGAAGAT
CGAGCCCCTGGGCGTGGCCCCCACCAAGGCCAAGCTGACCGTGCAGGCCCGCCAGCTGCTGTCCGGCATCGTGCAGC
AGCAGTCCAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTG
CAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGGAGATCTGGGACAACATGACCTGGATGGA
GTGGGAGCGCGAGATCAACAACTACACCGACATCATCTACTCCCTGATCGAGGAGTCCCAGAACCAGCAGGAGAAGA
ACGAGCAGGAGCTGCTGGCCCTGGACAAGTGGGCCTCCCTGTGGAACTGGTTCGACATCACCAACTGGCTGTGGGGC
CTGAACGACATCTTCGAGGCCCAGAAGATCGAGTGGCACGAGTAGGGATCCTCTAGA

>HV130111_N141A CON-Sgp140CFI_avi N141A
AAGCTTGTCGACACCATGCGCGTGCGCGGCATCCAGCGCAACTGCCAGCACCTGTGGCGCTGGGGCACCCTGATCCT
GGGCATGCTGATGATCTGCTCCGCCGCCGAGAACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGG
CCAACACCACCCTGTTCTGCGCCTCCGACGCCAAGGCCTACGACACCGAGGTGCACAACGTGTGGGCCACCCACGCC
TGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAA
CATGGTGGAGCAGATGCACGAGGACATCATCTCCCTGTGGGACCAGTCCCTGAAGCCCTGCGTGAAGCTGACCCCCC
TGTGCGTGACCCTGAACTGCACCAACGTGAACGTGACCAACACCACCGCCAACACCGAGGAGAAGGGCGAGATCAAG
AACTGCTCCTTCAACATCACCACCGAGATCCGCGACAAGAAGCAGAAGGTGTACGCCCTGTTCTACCGCCTGGACGT
GGTGCCCATCGACGACAACAACAACAACTCCTCCAACTACCGCCTGATCAACTGCAACACCTCCGCCATCACCCAGG
CCTGCCCCAAGGTGTCCTTCGAGCCCATCCCCATCCACTACTGCGCCCCCGCCGGCTTCGCCATCCTGAAGTGCAAC
GACAAGAAGTTCAACGGCACCGGCCCCTGCAAGAACGTGTCCACCGTGCAGTGCACCCACGGCATCAAGCCCGTGGT
GTCCACCCAGCTGCTGCTGAACGGCTCCCTGGCCGAGGAGGAGATCATCATCCGCTCCGAGAACATCACCAACAACG
CCAAGACCATCATCGTGCAGCTGAACGAGTCCGTGGAGATCAACTGCACCCGCCCCAACAACAACACCCGCAAGTCC
ATCCGCATCGGCCCCGGCCAGGCCTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACAT
CTCCGGCACCAAGTGGAACAAGACCCTGCAGCAGGTGGCCAAGAAGCTGCGCGAGCACTTCAACAACAAGACCATCA
TCTTCAAGCCCTCCTCCGGCGGCGACCTGGAGATCACCACCCACTCCTTCAACTGCCGCGGCGAGTTCTTCTACTGC
AACACCTCCGGCCTGTTCAACTCCACCTGGATCGGCAACGGCACCAAGAACAACAACAACACCAACGACACCATCAC
CCTGCCCTGCCGCATCAAGCAGATCATCAACATGTGGCAGGGCGTGGGCCAGGCCATGTACGCCCCCCCCATCGAGG
GCAAGATCACCTGCAAGTCCAACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAACAACAACACCAACGAGACC
GAGATCTTCCGCCCCGGCGGCGGCGACATGCGCGACAACTGGCGCTCCGAGCTGTACAAGTACAAGGTGGTGAAGAT
CGAGCCCCTGGGCGTGGCCCCCACCAAGGCCAAGCTGACCGTGCAGGCCCGCCAGCTGCTGTCCGGCATCGTGCAGC
AGCAGTCCAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTG
CAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGGAGATCTGGGACAACATGACCTGGATGGA
GTGGGAGCGCGAGATCAACAACTACACCGACATCATCTACTCCCTGATCGAGGAGTCCCAGAACCAGCAGGAGAAGA
ACGAGCAGGAGCTGCTGGCCCTGGACAAGTGGGCCTCCCTGTGGAACTGGTTCGACATCACCAACTGGCTGTGGGGC
CTGAACGACATCTTCGAGGCCCAGAAGATCGAGTGGCACGAGTAGGGATCCTCTAGA

>HV130111_V1_4Q CON-Sgp140CFI_avi V1_4Q
AAGCTTGTCGACACCATGCGCGTGCGCGGCATCCAGCGCAACTGCCAGCACCTGTGGCGCTGGGGCACCCTGATCCT
GGGCATGCTGATGATCTGCTCCGCCGCCGAGAACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGG
CCAACACCACCCTGTTCTGCGCCTCCGACGCCAAGGCCTACGACACCGAGGTGCACAACGTGTGGGCCACCCACGCC
TGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAA
CATGGTGGAGCAGATGCACGAGGACATCATCTCCCTGTGGGACCAGTCCCTGAAGCCCTGCGTGAAGCTGACCCCCC
TGTGCGTGACCCTGCAATGCACCAACGTGCAAGTGACCCAAACCACCCAAAACACCGAGGAGAAGGGCGAGATCAAG
AACTGCTCCTTCAACATCACCACCGAGATCCGCGACAAGAAGCAGAAGGTGTACGCCCTGTTCTACCGCCTGGACGT
GGTGCCCATCGACGACAACAACAACAACTCCTCCAACTACCGCCTGATCAACTGCAACACCTCCGCCATCACCCAGG
CCTGCCCCAAGGTGTCCTTCGAGCCCATCCCCATCCACTACTGCGCCCCCGCCGGCTTCGCCATCCTGAAGTGCAAC
GACAAGAAGTTCAACGGCACCGGCCCCTGCAAGAACGTGTCCACCGTGCAGTGCACCCACGGCATCAAGCCCGTGGT
GTCCACCCAGCTGCTGCTGAACGGCTCCCTGGCCGAGGAGGAGATCATCATCCGCTCCGAGAACATCACCAACAACG
CCAAGACCATCATCGTGCAGCTGAACGAGTCCGTGGAGATCAACTGCACCCGCCCCAACAACAACACCCGCAAGTCC
ATCCGCATCGGCCCCGGCCAGGCCTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACAT
CTCCGGCACCAAGTGGAACAAGACCCTGCAGCAGGTGGCCAAGAAGCTGCGCGAGCACTTCAACAACAAGACCATCA
TCTTCAAGCCCTCCTCCGGCGGCGACCTGGAGATCACCACCCACTCCTTCAACTGCCGCGGCGAGTTCTTCTACTGC

Figure 10 (cont)

AACACCTCCGGCCTGTTCAACTCCACCTGGATCGGCAACGGCACCAAGAACAACAACAACACCAACGACACCATCAC
CCTGCCCTGCCGCATCAAGCAGATCATCAACATGTGGCAGGGCGTGGGCCAGGCCATGTACGCCCCCCCCATCGAGG
GCAAGATCACCTGCAAGTCCAACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAACAACAACACCAACGAGACC
GAGATCTTCCGCCCCGGCGGCGGCGACATGCGCGACAACTGGCGCTCCGAGCTGTACAAGTACAAGGTGGTGAAGAT
CGAGCCCCTGGGCGTGGCCCCCACCAAGGCCAAGCTGACCGTGCAGGCCCGCCAGCTGCTGTCCGGCATCGTGCAGC
AGCAGTCCAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTG
CAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGGAGATCTGGGACAACATGACCTGGATGGA
GTGGGAGCGCGAGATCAACAACTACACCGACATCATCTACTCCCTGATCGAGGAGTCCCAGAACCAGCAGGAGAAGA
ACGAGCAGGAGCTGCTGGCCCTGGACAAGTGGGCCTCCCTGTGGAACTGGTTCGACATCACCAACTGGCTGTGGGGC
CTGAACGACATCTTCGAGGCCCAGAAGATCGAGTGGCACGAGTAGGGATCCTCTAGA

Translate results

>HV1301184
MRVRGIQRNCQHLWRWGTLILGMLMICSAAENLWVTVYYGVPVWKEANTTLFCA
SDAKAYDTEVHNVWATHACVPTDPNPQEIVLENVTENFNMWKNNMVEQMHEDIISLWDQS
LKPCVKLTPLCVTLNCTNVNVTNTTNNTEEKGEIKNCSFNITTEIRDKKQKVYALFYRLD
VVPIDDNNNNSSNYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGTG
PCKNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENITNNAKTIIVQLNESVEINCT
RPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNISGTKWNKTLQQVAKKLREHFNNKTI
IFKPSSGGDLEITTHSFNCRGEFFYCNTSGLFNSTWIGNGTKNNNNTNDTITLPCRIKQI
INMWQGVGQAMYAPPIEGKITCKSNITGLLLTRDGGNNNTNETEIFRPGGGDMRDNWRSE
LYKYKVVKIEPLGVAPTKCKERVVGRRRRRRAVGIGAVFLGFLGAAGSTMGAASITLTVQ
ARQLLSGIVQQQSNLLRAPEAQQHLLQLTVWGIKQLQARVLAVERYLKDQQLLGIWGCSG
KLICCTTVPWNSSWSNKSQDEIWDNMTWMEWEREINNYTDIIYSLIEESQNQQEKNEQEL
LALD*

>HV1301185
MRVRGIQRNCQHLWRWGTLILGMLMICSAAENLWVTVYYGVPVWKEANTTLFCA
SDAKAYDTEVHNVWATHACVPTDPNPQEIVLENVTENFNMWKNNMVEQMHEDIISLWDQS
LKPCVKLTPLCVTLNCTNVNVTNTTNNTEEKGEIKNCSFNITTEIRDKKQKVYALFYRLD
VVPIDDNNNNSSNYRLINCNTSACTQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGTG
PCKNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENITNNAKTIIVQLNESVEINCT
RPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNISGTKWNKTLQQVAKKLREHFNNKTI
IFKPSSGGDLEITTHSFNCRGEFFYCNTSGLFNSTWIGNGTKNNNNTNDTITLPCRIKQI
INMWQGVGQCMYAPPIEGKITCKSNITGLLLTRDGGNNNTNETEIFRPGGGDMRDNWRSE
LYKYKVVKIEPLGVAPTKCKERVVGRRRRRRAVGIGAVFLGFLGAAGSTMGAASITLTVQ
ARQLLSGIVQQQSNLLRAPEAQQHLLQLTVWGIKQLQARVLAVERYLKDQQLLGIWGCSG
KLICCTTVPWNSSWSNKSQDEIWDNMTWMEWEREINNYTDIIYSLIEESQNQQEKNEQEL
LALD*

>HV1301186
MRVRGIQRNCQHLWRWGTLILGMLMICSAAENLWVTVYYGVPVWKEANTTLFCA
SDAKAYDTEVHNVWATHACVPTDPNPQEIVLENVTENFNMWKNNMVEQMHEDIISLWDQS
LKPCVKLTPLCVTLNCTNVNVTNTTNNTEEKGEIKNCSFNITTEIRDKKQKVYALFYRLD
VVPIDDNNNNSSNYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGTG
PCKNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENITNNAKTIIVQLNESVEINCT
RPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNISGTKWNKTLQQVAKKLREHFNNKTI
IFKPSSGGDLEITTHSFNCRGEFFYCNTSGLFNSTWIGNGTKNNNNTNDTITLPCRIKQI
INMWQGVGQAMYAPPIEGKITCKSNITGLLLTRDGGNNNTNETEIFRPGGGDMRDNWRSE
LYKYKVVKIEPLGVAPTKCKERVVGRRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQ
ARQLLSGIVQQQSNLLRAPEAQQHLLQLTVWGIKQLQARVLAVERYLKDQQLLGIWGCSG
KLICCTTVPWNSSWSNKSQDEIWDNMTWMEWEREINNYTDIIYSLIEESQNQQEKNEQEL
LALD*

Figure 10 (cont)

>HV1301187
MRVRGIQRNCQHLWRWGTLILGMLMICSAAENLWVTVYYGVPVWKEANTTLFCA
SDAKAYDTKVHNVWATHACVPTDPNPQEIVLENVTENFNMWKNNMVEQMHEDIISLWDQS
LKPCVKLTPLCVTLNCTNVNVTNTTNNTEEKGEIKNCSFNITTEIRDKKQKVYALFYRLD
VVPIDDNNNNSSNYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGTG
PCKNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENITNNAKTIIVQLNESVEINCT
RPNNNTRKSIRIGPGQWFYATGDIIGDIRQAHCNISGTKWNKTLQQVAKKLREHFNNKTI
IFKPSSGGDLEITTHSFNCRGEFFYCNTSGLFNSTWIGNGTKNNNNTNDTITLPCRIKQI
INMWQGVGQAMYAPPIEGKITCKSNITGLLLTRDGGNNNTNETEIFRPGGGDMRDNWRSE
LYKYKVVKIEPLGVAPTKCKERVVGRRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQ
ARQLLSGIVQQQSNLLRAPEAQQHLLQLTVWGIKQLQARVLAVERYLKDQQLLGIWGCSG
KLICCTTVPWNSSWSNKSQDEIWDNMTWMEWEREINNYTDIIYSLIEESQNQQEKNEQEL
LALD*

>HV1301188
MRVRGIQRNCQHLWRWGTLILGMLMICSAAENLWVTVYYGVPVWKEANTTLFCA
SDAKAYDTEVRNVWATHACVPTDPNPQEIVLENVTENFNMWKNNMVEQMHEDIISLWDQS
LKPCVKLTPLCVTLNCTNVNVTNTTNNTEEKGEIKNCSFNITTEIRDKKQKVYALFYRLD
VVPIDDNNNNSSNYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGTG
PCKNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENITNNAKTIIVQLNESVEINCT
RPNNNTRKSIRIGPGQWFYATGDIIGDIRQAHCNISGTKWNKTLQQVAKKLREHFNNKTI
IFKPSSGGDLEITTHSFNCRGEFFYCNTSGLFNSTWIGNGTKNNNNTNDTITLPCRIKQI
INMWQGVGQAMYAPPIEGKITCKSNITGLLLTRDGGNNNTNETEIFRPGGGDMRDNWRSE
LYKYKVVKIEPLGVAPTKCKERVVGRRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQ
ARQLLSGIVQQQSNLLRAPEAQQHLLQLTVWGIKQLQARVLAVERYLKDQQLLGIWGCSG
KLICCTTVPWNSSWSNKSQDEIWDNMTWMEWEREINNYTDIIYSLIEESQNQQEKNEQEL
LALD*

>HV1301257
VDATMGSLQPLATLYLLGMLVASVLAAENLWVTVYYGVPVWKEANTTLFCASDAKAYDTE
VHNVWATHACVPTDPNPQEIVLENVTENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTP
LCVTLNCTNVNVTNTTNNTEEKGEIKNCSFNITTEIRDKKQKVYALFYRLDVVPIDDNNN
NSSNYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGTGPCKNVSTVQ
CTHGIKPVVSTQLLLNGSLAEEEIIIRSENITNNAKTIIVQLNESVEINCTRPNNNTRKS
IRIGPGQAFYATGDIIGDIRQAHCNISGTKWNKTLQQVAKKLREHFNNKTIIFKPSSGGD
LEITTHSFNCRGEFFYCNTSGLFNSTWIGNGTKNNNNTNDTITLPCRIKQIINMWQGVGQ
AMYAPPIEGKITCKSNITGLLLTRDGGNNNTNETEIFRPGGGDMRDNWRSELYKYKVVKI
EPLGVAPTRCKRRVVGRRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIV
QQQSNLLRAPEAQQHLLKLTVWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNVP
WNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALDGSGLN
DIFEAQKIEWHE*

>HV1301258
VDATMGSLQPLATLYLLGMLVASVLAAENLWVTVYYGVPVWKEANTTLFCASDAKAYDTE
VHNVWATHACVPTDPNPQEIVLENVTENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTP
LCVTLNCTNVNVTNTTNNTEEKGEIKNCSFNITTEIRDKKQKVYALFYRLDVVPIDDNNN
NSSNYRLINCNTSACTQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGTGPCKNVSTVQ
CTHGIKPVVSTQLLLNGSLAEEEIIIRSENITNNAKTIIVQLNESVEINCTRPNNNTRKS
IRIGPGQAFYATGDIIGDIRQAHCNISGTKWNKTLQQVAKKLREHFNNKTIIFKPSSGGD
LEITTHSFNCRGEFFYCNTSGLFNSTWIGNGTKNNNNTNDTITLPCRIKQIINMWQGVGQ
CMYAPPIEGKITCKSNITGLLLTRDGGNNNTNETEIFRPGGGDMRDNWRSELYKYKVVKI
EPLGVAPTRCKRRVVGRRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIV
QQQSNLLRAPEAQQHLLKLTVWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNVP
WNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALDGSGLN
DIFEAQKIEWHE*

Figure 10 (cont)

>HV1301259
VDATMGSLQPLATLYLLGMLVASVLAAENLWVTVYYGVPVWKEANTTLFCASDAKAYDTK
VHNVWATHACVPTDPNPQEIVLENVTENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTP
LCVTLNCTNVNVTNTTNNTEEKGEIKNCSFNITTEIRDKKQKVYALFYRLDVVPIDDNNN
NSSNYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGTGPCKNVSTVQ
CTHGIKPVVSTQLLLNGSLAEEEIIIRSENITNNAKTIIVQLNESVEINCTRPNNNTRKS
IRIGPGQWFYATGDIIGDIRQAHCNISGTKWNKTLQQVAKKLREHFNNKTIIFKPSSGGD
LEITTHSFNCRGEFFYCNTSGLFNSTWIGNGTKNNNNTNDTITLPCRIKQIINMWQGVGQ
AMYAPPIEGKITCKSNITGLLLTRDGGNNNTNETEIFRPGGGDMRDNWRSELYKYKVVKI
EPLGVAPTRCKRRVVGRRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIV
QQQSNLLRAPEAQQHLLKLTVWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNVP
WNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALDGSGLN
DIFEAQKIEWHE*

>HV1301260
VDATMGSLQPLATLYLLGMLVASVLAAENLWVTVYYGVPVWKEANTTLFCASDAKAYDTE
VRNVWATHACVPTDPNPQEIVLENVTENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTP
LCVTLNCTNVNVTNTTNNTEEKGEIKNCSFNITTEIRDKKQKVYALFYRLDVVPIDDNNN
NSSNYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGTGPCKNVSTVQ
CTHGIKPVVSTQLLLNGSLAEEEIIIRSENITNNAKTIIVQLNESVEINCTRPNNNTRKS
IRIGPGQWFYATGDIIGDIRQAHCNISGTKWNKTLQQVAKKLREHFNNKTIIFKPSSGGD
LEITTHSFNCRGEFFYCNTSGLFNSTWIGNGTKNNNNTNDTITLPCRIKQIINMWQGVGQ
AMYAPPIEGKITCKSNITGLLLTRDGGNNNTNETEIFRPGGGDMRDNWRSELYKYKVVKI
EPLGVAPTRCKRRVVGRRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIV
QQQSNLLRAPEAQQHLLKLTVWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNVP
WNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALDGSGLN
DIFEAQKIEWHE*

>HV1301521_ferritin
*VTVVDAST*MGSLQPLATLYLLGMLVASVLAAENLWVTVYYGVPVWKEANTTLFCASDAKA
YDTEVHNVWATHACVPTDPNPQEIVLENVTENFNMWKNNMVEQMHEDIISLWDQSLKPCV
KLTPLCVTLDCTNVKVTSTTSNTEEKGEIKNCSFNITTEIRDKKQKVYALFYRLDVVPID
DNNNNSSNYRLINCNTSACTQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGTGPCKNV
STVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENITNNAKTIIVQLNESVEINCTRPNNN
TRKSIRIGPGQAFYATGDIIGDIRQAHCNISGTKWNKTLQQVAKKLREHFNNKTIIFKPS
SGGDLEITTHSFNCRGEFFYCNTSGLFNSTWIGNGTKNNNNTNDTITLPCRIKQIINMWQ
GVGQCMYAPPIEGKITCKSNITGLLLTRDGGNNNTNETEIFRPGGGDMRDNWRSELYKYK
VVKIEPLGVAPTRCKRRVVGRRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLL
SGIVQQQSNLLRAPEAQQHLLKLTVWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICC
TNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALD**G**
GGSGDIIKLLNEQVNKEMNSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIF
LNENNVPVQLTSISAPEHKFEGLTQIFQKAYEHEQHISESINNIVDHAIKSKDHATFNFL
QWYVAEQHEEEVLFKDILDKIELIGNENHGLYLADQYVKGIAKSRKS**

>HV1301521(originally HV1301405_N130D_N135K_N138S_N141S)
VDASTMGSLQPLATLYLLGMLVASVLAAENLWVTVYYGVPVWKEANTTLFCASDAKAYDT
EVHNVWATHACVPTDPNPQEIVLENVTENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLT
PLCVTLDCTNVKVTSTTSNTEEKGEIKNCSFNITTEIRDKKQKVYALFYRLDVVPIDDNN
NNSSNYRLINCNTSACTQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGTGPCKNVSTV
QCTHGIKPVVSTQLLLNGSLAEEEIIIRSENITNNAKTIIVQLNESVEINCTRPNNNTRK
SIRIGPGQAFYATGDIIGDIRQAHCNISGTKWNKTLQQVAKKLREHFNNKTIIFKPSSGG
DLEITTHSFNCRGEFFYCNTSGLFNSTWIGNGTKNNNNTNDTITLPCRIKQIINMWQGVG
QCMYAPPIEGKITCKSNITGLLLTRDGGNNNTNETEIFRPGGGDMRDNWRSELYKYKVVK
IEPLGVAPTRCKRRVVGRRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGI

Figure 10 (cont)

VQQQSNLLRAPEAQQHLLKLTVWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNV
PWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALD*
GS

>HV1301405_N138S_N141S
VDASTMGSLQPLATLYLLGMLVASVLAAENLWVTVYYGVPVWKEANTTLFCASDAKAYDT
EVHNVWATHACVPTDPNPQEIVLENVTENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLT
PLCVTLNCTNVNVTSTTSNTEEKGEIKNCSFNITTEIRDKKQKVYALFYRLDVVPIDDNN
NNSSNYRLINCNTSACTQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGTGPCKNVSTV
QCTHGIKPVVSTQLLLNGSLAEEEIIIRSENITNNAKTIIVQLNESVEINCTRPNNNTRK
SIRIGPGQAFYATGDIIGDIRQAHCNISGTKWNKTLQQVAKKLREHFNNKTIIFKPSSGG
DLEITTHSFNCRGEFFYCNTSGLFNSTWIGNGTKNNNNTNDTITLPCRIKQIINMWQGVG
QCMYAPPIEGKITCKSNITGLLLTRDGGNNNTNETEIFRPGGGDMRDNWRSELYKYKVVK
IEPLGVAPTRCKRRVVGRRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGI
VQQQSNLLRAPEAQQHLLKLTVWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNV
PWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALD*
GS

>HV1301405
VDASTMGSLQPLATLYLLGMLVASVLAAENLWVTVYYGVPVWKEANTTLFCASDAKAYDT
EVHNVWATHACVPTDPNPQEIVLENVTENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLT
PLCVTLNCTNVNVTNTTNNTEEKGEIKNCSFNITTEIRDKKQKVYALFYRLDVVPIDDNN
NNSSNYRLINCNTSACTQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGTGPCKNVSTV
QCTHGIKPVVSTQLLLNGSLAEEEIIIRSENITNNAKTIIVQLNESVEINCTRPNNNTRK
SIRIGPGQAFYATGDIIGDIRQAHCNISGTKWNKTLQQVAKKLREHFNNKTIIFKPSSGG
DLEITTHSFNCRGEFFYCNTSGLFNSTWIGNGTKNNNNTNDTITLPCRIKQIINMWQGVG
QCMYAPPIEGKITCKSNITGLLLTRDGGNNNTNETEIFRPGGGDMRDNWRSELYKYKVVK
IEPLGVAPTRCKRRVVGRRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGI
VQQQSNLLRAPEAQQHLLKLTVWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNV
PWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALD*
GS

>HV1301258_N301A
VDATMGSLQPLATLYLLGMLVASVLAAENLWVTVYYGVPVWKEANTTLFCASDAKAYDTE
VHNVWATHACVPTDPNPQEIVLENVTENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTP
LCVTLNCTNVNVTNTTNNTEEKGEIKNCSFNITTEIRDKKQKVYALFYRLDVVPIDDNNN
NSSNYRLINCNTSACTQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGTGPCKNVSTVQ
CTHGIKPVVSTQLLLNGSLAEEEIIIRSENITNNAKTIIVQLNESVEINCTRPNANTRKS
IRIGPGQAFYATGDIIGDIRQAHCNISGTKWNKTLQQVAKKLREHFNNKTIIFKPSSGGD
LEITTHSFNCRGEFFYCNTSGLFNSTWIGNGTKNNNNTNDTITLPCRIKQIINMWQGVGQ
CMYAPPIEGKITCKSNITGLLLTRDGGNNNTNETEIFRPGGGDMRDNWRSELYKYKVVKI
EPLGVAPTRCKRRVVGRRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIV
QQQSNLLRAPEAQQHLLKLTVWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNVP
WNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALDGSGLN
DIFEAQKIEWHE*

>HV1301258_N332A, CON-Schim.6R.DS.SOSIP.664_avi_N332A
VDATMGSLQPLATLYLLGMLVASVLAAENLWVTVYYGVPVWKEANTTLFCASDAKAYDTE
VHNVWATHACVPTDPNPQEIVLENVTENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTP
LCVTLNCTNVNVTNTTNNTEEKGEIKNCSFNITTEIRDKKQKVYALFYRLDVVPIDDNNN
NSSNYRLINCNTSACTQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGTGPCKNVSTVQ
CTHGIKPVVSTQLLLNGSLAEEEIIIRSENITNNAKTIIVQLNESVEINCTRPNNNTRKS
IRIGPGQAFYATGDIIGDIRQAHCAISGTKWNKTLQQVAKKLREHFNNKTIIFKPSSGGD
LEITTHSFNCRGEFFYCNTSGLFNSTWIGNGTKNNNNTNDTITLPCRIKQIINMWQGVGQ

Figure 10 (cont)

CMYAPPIEGKITCKSNITGLLLTRDGGNNNTNETEIFRPGGGDMRDNWRSELYKYKVVKI
EPLGVAPTRCKRRVVGRRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIV
QQQSNLLRAPEAQQHLLKLTVWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNVP
WNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALDGSGLN
DIFEAQKIEWHE*

>HV1301521_c_sorta
VDASTMGSLQPLATLYLLGMLVASVLAAENLWVTVYYGVPVWKEANTTLFCASDAKAYDT
EVHNVWATHACVPTDPNPQEIVLENVTENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLT
PLCVTLDCTNVKVTSTTSNTEEKGEIKNCSFNITTEIRDKKQKVYALFYRLDVVPIDDNN
NNSSNYRLINCNTSACTQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGTGPCKNVSTV
QCTHGIKPVVSTQLLLNGSLAEEEIIIRSENITNNAKTIIVQLNESVEINCTRPNNNTRK
SIRIGPGQAFYATGDIIGDIRQAHCNISGTKWNKTLQQVAKKLREHFNNKTIIFKPSSGG
DLEITTHSFNCRGEFFYCNTSGLFNSTWIGNGTKNNNNTNDTITLPCRIKQIINMWQGVG
QCMYAPPIEGKITCKSNITGLLLTRDGGNNNTNETEIFRPGGGDMRDNWRSELYKYKVVK
IEPLGVAPTRCKRRVVGRRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGI
VQQQSNLLRAPEAQQHLLKLTVWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNV
PWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALDLPST
GG**

>HV1301639_avi
MGSLQPLATLYLLGMLVASVLAAENLWVTVYYGVPVWKEANTTLFCASDAKAYDTEVHNV
WATHACVPTDPNPQEIVLENVTENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVT
LDCTNVKVTNTTNNTEEKGEIKNCSFNITTEIRDKKQKVYALFYRLDVVPIDDNNNNSSN
YRLINCNTSACTQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGTGPCKNVSTVQCTHG
IKPVVSTQLLLNGSLAEEEIIIRSENITNNAKTIIVQLNESVEINCTRPNNNTRKSIRIG
PGQAFYATGDIIGDIRQAHCNISGTKWNKTLQQVAKKLREHFNNKTIIFKPSSGGDLEIT
THSFNCRGEFFYCNTSGLFNSTWIGNGTKNNNNTNDTITLPCRIKQIINMWQGVGQCMYA
PPIEGKITCKSNITGLLLTRDGGNNNTNETEIFRPGGGDMRDNWRSELYKYKVVKIEPLG
VAPTRCKRRVVGRRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQS
NLLRAPEAQQHLLKLTVWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNVPWNSS
WSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALDGSGLNDIFE
AQKIEWHE*

>HV1301640_avi
MGSLQPLATLYLLGMLVASVLAAENLWVTVYYGVPVWKEANTTLFCASDAKAYDTEVHNV
WATHACVPTDPNPQEIVLENVTENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVT
LNCTNVNVTSTTSNTEEKGEIKNCSFNITTEIRDKKQKVYALFYRLDVVPIDDNNNNSSN
YRLINCNTSACTQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGTGPCKNVSTVQCTHG
IKPVVSTQLLLNGSLAEEEIIIRSENITNNAKTIIVQLNESVEINCTRPNNNTRKSIRIG
PGQAFYATGDIIGDIRQAHCNISGTKWNKTLQQVAKKLREHFNNKTIIFKPSSGGDLEIT
THSFNCRGEFFYCNTSGLFNSTWIGNGTKNNNNTNDTITLPCRIKQIINMWQGVGQCMYA
PPIEGKITCKSNITGLLLTRDGGNNNTNETEIFRPGGGDMRDNWRSELYKYKVVKIEPLG
VAPTRCKRRVVGRRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQS
NLLRAPEAQQHLLKLTVWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNVPWNSS
WSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALDGSGLNDIFE
AQKIEWHE*

>HV1301641_avi
MGSLQPLATLYLLGMLVASVLAAENLWVTVYYGVPVWKEANTTLFCASDAKAYDTEVHNV
WATHACVPTDPNPQEIVLENVTENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVT
LDCTNVKVTSTTSNTEEKGEIKNCSFNITTEIRDKKQKVYALFYRLDVVPIDDNNNNSSN
YRLINCNTSACTQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGTGPCKNVSTVQCTHG
IKPVVSTQLLLNGSLAEEEIIIRSENITNNAKTIIVQLNESVEINCTRPNNNTRKSIRIG

Figure 10 (cont)

PGQAFYATGDIIGDIRQAHCNISGTKWNKTLQQVAKKLREHFNNKTIIFKPSSGGDLEIT
THSFNCRGEFFYCNTSGLFNSTWIGNGTKNNNNTNDTITLPCRIKQIINMWQGVGQCMYA
PPIEGKITCKSNITGLLLTRDGGNNNTNETEIFRPGGGDMRDNWRSELYKYKVVKIEPLG
VAPTRCKRRVVGRRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQS
NLLRAPEAQQHLLKLTVWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNVPWNSS
WSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALDGSGLNDIFE
AQKIEWHE*

>HV1301613
MGSLQPLATLYLLGMLVASVLAAENLWVTVYYGVPVWKEANTTLFCASDAKAYDTKVHNV
WATHACVPTDPNPQEIVLENVTENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVT
LNCTNVNVTNTTNNTEEKGEIKNCSFNITTEIRDKKQKVYALFYRLDVVPIDDNNNNSSN
YRLINCNTSACTQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGTGPCKNVSTVQCTHG
IKPVVSTQLLLNGSLAEEEIIIRSENITNNAKTIIVQLNESVEINCTRPNNNTRKSIRIG
PGQWFYATGDIIGDIRQAHCNISGTKWNKTLQQVAKKLREHFNNKTIIFKPSSGGDLEIT
THSFNCRGEFFYCNTSGLFNSTWIGNGTKNNNNTNDTITLPCRIKQIINMWQGVGQCMYA
PPIEGKITCKSNITGLLLTRDGGNNNTNETEIFRPGGGDMRDNWRSELYKYKVVKIEPLG
VAPTRCKRRVVGRRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQS
NLLRAPEAQQHLLKLTVWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNVPWNSS
WSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALD**

>HV130111_N137A CON-Sgp140CFI_avi N137A
KLVDTMRVRGIQRNCQHLWRWGTLILGMLMICSAAENLWVTVYYGVPVWKEANTTLFCAS
DAKAYDTEVHNVWATHACVPTDPNPQEIVLENVTENFNMWKNNMVEQMHEDIISLWDQSL
KPCVKLTPLCVTLNCTNVNVTATTNNTEEKGEIKNCSFNITTEIRDKKQKVYALFYRLDV
VPIDDNNNNSSNYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGTGP
CKNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENITNNAKTIIVQLNESVEINCTR
PNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNISGTKWNKTLQQVAKKLREHFNNKTII
FKPSSGGDLEITTHSFNCRGEFFYCNTSGLFNSTWIGNGTKNNNNTNDTITLPCRIKQII
NMWQGVGQAMYAPPIEGKITCKSNITGLLLTRDGGNNNTNETEIFRPGGGDMRDNWRSEL
YKYKVVKIEPLGVAPTKAKLTVQARQLLSGIVQQQSNLLRAIEAQQHLLQLTVWGIKQLQ
ARVLAVERYLKDQQLEIWDNMTWMEWEREINNYTDIIYSLIEESQNQQEKNEQELLALDK
WASLWNWFDITNWLWGLNDIFEAQKIEWHE*

>HV130111_N141A CON-Sgp140CFI_avi N141A
KLVDTMRVRGIQRNCQHLWRWGTLILGMLMICSAAENLWVTVYYGVPVWKEANTTLFCAS
DAKAYDTEVHNVWATHACVPTDPNPQEIVLENVTENFNMWKNNMVEQMHEDIISLWDQSL
KPCVKLTPLCVTLNCTNVNVTNTTANTEEKGEIKNCSFNITTEIRDKKQKVYALFYRLDV
VPIDDNNNNSSNYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGTGP
CKNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENITNNAKTIIVQLNESVEINCTR
PNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNISGTKWNKTLQQVAKKLREHFNNKTII
FKPSSGGDLEITTHSFNCRGEFFYCNTSGLFNSTWIGNGTKNNNNTNDTITLPCRIKQII
NMWQGVGQAMYAPPIEGKITCKSNITGLLLTRDGGNNNTNETEIFRPGGGDMRDNWRSEL
YKYKVVKIEPLGVAPTKAKLTVQARQLLSGIVQQQSNLLRAIEAQQHLLQLTVWGIKQLQ
ARVLAVERYLKDQQLEIWDNMTWMEWEREINNYTDIIYSLIEESQNQQEKNEQELLALDK
WASLWNWFDITNWLWGLNDIFEAQKIEWHE*

>HV130111_V1_4Q CON-Sgp140CFI_avi V1_4Q
KLVDTMRVRGIQRNCQHLWRWGTLILGMLMICSAAENLWVTVYYGVPVWKEANTTLFCAS
DAKAYDTEVHNVWATHACVPTDPNPQEIVLENVTENFNMWKNNMVEQMHEDIISLWDQSL
KPCVKLTPLCVTLQCTNVQVTQTTQNTEEKGEIKNCSFNITTEIRDKKQKVYALFYRLDV
VPIDDNNNNSSNYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGTGP
CKNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENITNNAKTIIVQLNESVEINCTR
PNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNISGTKWNKTLQQVAKKLREHFNNKTII

Figure 10 (cont)

FKPSSGGDLEITTHSFNCRGEFFYCNTSGLFNSTWIGNGTKNNNNTNDTITLPCRIKQII
NMWQGVGQAMYAPPIEGKITCKSNITGLLLTRDGGNNNTNETEIFRPGGGDMRDNWRSEL
YKYKVVKIEPLGVAPTKAKLTVQARQLLSGIVQQQSNLLRAIEAQQHLLQLTVWGIKQLQ
ARVLAVERYLKDQQLEIWDNMTWMEWEREINNYTDIIYSLIEESQNQQEKNEQELLALDK
WASLWNWFDITNWLWGLNDIFEAQKIEWHE*

Figure 10 (cont)

>HV1301405_D1305V1;CON-Schim.6R.DS.SOSIP.664_OPT_D1305V1

MGSLQPLATLYLLGMLVASVLAAENLWVTVYYGVPVWKEANTTLFCASDAKAYDTEVHNVWATHACVPTD
PNPQEIVLENVTENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPL**CVTLNCSTATVNNRAVDEMKNC**S
FNITTEIRDKKQKVYALFYRLDVVPIDDNNNNSSNYRLINCNTSACTQACPKVSFEPIPIHYCAPAGFAI
LKCNDKKFNGTGPCKNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENITNNAKTIIVQLNESVEIN
CTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNISGTKWNKTLQQVAKKLREHFNNKTIIFKPSSGG
DLEITTHSFNCRGEFFYCNTSGLFNSTWIGNGTKNNNNTNDTITLPCRIKQIINMWQGVGQCMYAPPIEG
KITCKSNITGLLLTRDGGNNNTNETEIFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTRCKRRVVGRRR
RRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQQHLLKLTVWGIKQLQA
RVLAVERYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEE
SQNQQEKNEQDLLALD**

Figure 11A

>HV1301521_ferritin
VTVVDAST<u>MGSLQPLATLYLLGMLVASVLA</u>AENLWVTVYYGVPVWKEANTTLFCASDAKA
YDTEVHNVWATHACVPTDPNPQEIVLENVTENFNMWKNNMVEQMHEDIISLWDQSLKPCV
KLTPLCVTLDCTNVKVTSTTSNTEEKGEIKNCSFNITTEIRDKKQKVYALFYRLDVVPID
DNNNNSSNYRLINCNTSACTQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGTGPCKNV
STVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENITNNAKTIIVQLNESVEINCTRPNNN
TRKSIRIGPGQAFYATGDIIGDIRQAHCNISGTKWNKTLQQVAKKLREHFNNKTIIFKPS
SGGDLEITTHSFNCRGEFFYCNTSGLFNSTWIGNGTKNNNNTNDTITLPCRIKQIINMWQ
GVGQCMYAPPIEGKITCKSNITGLLLTRDGGNNNTNETEIFRPGGGDMRDNWRSELYKYK
VVKIEPLGVAPTRCKRRVVGRRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLL
SGIVQQQSNLLRAPEAQQHLLKLTVWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICC
TNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALD**G**
GGSGDIIKLLNEQVNKEMNSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIF
LNENNVPVQLTSISAPEHKFEGLTQIFQKAYEHEQHISESINNIVDHAIKSKDHATFNFL
QWYVAEQHEEEVLFKDILDKIELIGNENHGLYLADQYVKGIAKSRKS

Italicized is cloning sites

<u>Underlined</u> is position of signal peptide of the ConS envelope

GGGS in bold is one embodiment of linker between ConS and ferritin protein

Figure 11B

>HV1301580_D230N_H289N_P291S;CH848.3.D1305.10.19_D949V3.DS.SOSIP_D230N_H289N_P291S (glycan hole filled)

ATGGGATCTCTGCAACCTCTGGCCACACTGTACCTGCTGGGAATGCTGGTGGCTTCTGTGCTGGCCGCCG
AGAATCTGTGGGTCACAGTGTACTATGGCGTGCCCGTGTGGAAAGAGGCCAAGACCACACTGTTCTGTGC
CAGCGACGCCAAGGCCTACAAGAAAGAGGTGCACAACGTCTGGGCCACACACGCCTGTGTGCCTACCGAT
CCATCTCCTCAAGAGCTGTTCCTGGAAAACGTGACCGAGAACTTCAACATGTGGAAGAACGACATGGTGG
ACCAGATGCACGAGGACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCTTGCGTGAAGCTGACCCCTCT
GTGCGTGACCCTGATCTGTAGCACCGCCACCGTGAACAACAGAGCCGTGGACGAGATGAAGAACTGCAGC
TTCAACACCACCACCGAGATCCGGGACAAGAAGAAGAAAGAGTACGCCCTGTTCTATCGGAGCGACGTGG
TGCCCCTGGACGAGACAAACAACACCAGCGAGTACCGGCTGATCAACTGCAACACCTCCGCCTGCACTCA
GGCCTGTCCTAAAGTGACCTTCGAGCCCATTCCTATCCACTACTGTGCCCCTGCCGGCTACGCCATCCTG
AAGTGCAACNACGAGACATTCAACGGCACAGGCCCCTGCAGCAATGTGTCCACCGTGCAGTGTACCCACG
GCATCAGACCAGTGGTGTCTACCCAGCTGCTGCTGAATGGAAGCCTGGCCGAGAAAGAAATCGTGATCAG
AAGCGAGAACCTGACCAACAACGCCAAGATCATCATTGTGCATCTGNACACCNCTGTGGAAATCGTGTGC
ACCCGGCCTAACAACAACACCCGGAAGTCTGTGCGGATCGGCCCTGGCCAGACATTCTATGCCACCGGCG
ATATCATCGGCGACATCAAGCAGGCCCACTGCAACATCAGCGAGGAAAAGTGGAACGAGACACTGCAGAA
AGTGGGCATCGAGCTGCAGAAGCACTTCCCCAACAAGACCATCAAGTACAACCAGAGCGCTGGCGGCGAC
ATGGAAATCACCACACACAGCTTCAATTGTGGCGGCGAGTTCTTCTACTGCAATACCAGCAAGCTGTTCA
ACAGCACCTACAACGGCACCTATATCAGCACCAACTCCACCAATAGCACCAGCTACATCACCCTGCAGTG
CCGGATCAAGCAGATCATCAATATGTGGCAAGGCGTCGGCCGGTGTATGTACGCCCCTCCTATCGCCGGC
AACATCACCTGTCGGAGCAATATCACAGGCCTGCTGCTCACCAGAGATGGCGGCATCAACAACGTGTCCA
ACGAGACAGAAACCTTCCGGCCTGCCGGCGGAGACATGAGAGACAATTGGAGAAGCGAGCTGTACAAGTA
CAAGGTGGTCAAGATCGAGCCCCTGGGCGTCGCACCAACACGGTGCAAGAGAAGAGTCGTGGGCCGTCGT
AGAAGGCGGAGAGCCGTTGGAATTGGCGCCGTGTTCCTGGGCTTTCTGGGAGCCGCTGGATCTACAATGG
GCGCTGCCAGCATGACCCTGACAGTGCAGGCTAGAAATCTGCTGAGCGGCATCGTGCAGCAGCAGAGCAA
TCTGCTCAGAGCCCCTGAGGCTCAGCAGCACCTCCTGAAACTGACAGTGTGGGGAATCAAGCAGCTGCAG
GCCAGAGTGCTGGCAGTGGAAAGATACCTGAGGGACCAGCAGCTCCTCGGAATCTGGGGATGTAGCGGCA
AGCTGATCTGCTGCACCAACGTGCCCTGGAACTCCAGCTGGTCCAACCGGAATCTGAGCGAGATCTGGGA
TAACATGACCTGGCTGCAGTGGGACAAAGAGATCAGCAACTACACCCAGATCATCTACGGCCTGCTGGAA
GAGAGCCAGAACCAGCAAGAGAAAAACGAGCAGGACCTGCTGGCCCTGGACTGATAA

>HV1301502_D1305V1;JRFL_SOSIPv6_V1_PNGS_D1305V1 (V1 loop from 10.19)

ATGGGATCTCTGCAGCCTCTGGCCACACTGTACCTGCTGGGAATGCTGGTGGCCTCTTGTCTGGGCGTTG
AGAAGCTGTGGGTCACCGTGTACTATGGCGTGCCCGTGTGGAAAGAAGCCTGCACCACACTGTTCTGTGC
CAGCGACGCCAAGGCCTACGATACCAAGGTGCGCAATGTGTGGGCCACTCACTGCTGCGTGCCCACCGAT
CCTAATCCTCAAGAGGTGGTGCTGGAAAACGTGACCGAGCACTTCAACATGTGGAAGAACAACATGGTCG
AGCAGATGCAAGAGGACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCTTGCGTGAAGCTGACCCCTCT
GTGCGTGACCCTGAACTG*TAGCACCGCCACCGTGAACAACAGAGCCGTGGACGAGATG*AAGAACTGCAGC
TTCAACATCACCACCTCCATCAGAGACAAGGTGCAGAAAGAGTACGCCCTGTTCTACAAGCTGGACGTGG

Figure 12

TGCCCATCGACAACAACAACACCAGCTACAGACTGATCAGCTGCGACACCAGCGTGATCACCCAGGCCTG
TCCTAAGATCAGCTTCGAGCCCATTCCTATCCACTACTGTGCCCCTGCCGGCTTCGCCATCCTGAAGTGC
AACGACAAGACCTTCAACGGCAAGGGCCCCTGCAAGAACGTGTCCACCGTGCAGTGTACCCACGGCATCA
GACCAGTGGTGTCTACCCAGCTGCTGCTGAATGGCTCTCTGGCCGAGGAAGAAGTGGTCATCAGAAGCGA
CAACTTCACCAACAACGCCAAGACCATCATCGTGCAGCTGAAAGAGAGCGTCGAGATCAACTGCACCCGG
CCTAACAACAATACCCGGAAGTCCATCCACATCGGCCCTGGCAGATGGTTTTACACCACCGGCGAGATTA
TCGGCGACATCAGACAGGCCCACTGCAACATCAGCCGGGCCAAGTGGAACGACACCCTGAAGCAGATCGT
GATCAAGCTGAGAGAGCAGTTCGAGAACAAGACGATCGTGTTCAACCACAGCTCTGGCGGCGACCCCGAG
ATTGTGATGCACTCCTTTAACTGTGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACTCCA
CCTGGAACAACAACACAGAGGGCAGCAACAATACCGAGGGCAACACCATCACACTGCCCTGCCGGATCAA
GCAGATCATCAATATGTGGCAAGAGATCGGCAAGGCTATGTACGCCCCTCCTATCCGGGGCCAGATCAGA
TGCAGCAGCAATATCACAGGCCTGCTGCTCACCAGAGATGGCGGCATCAACGAGAACGGCACCGAGATCT
TCAGACCCGGCGGAGGCGACATGCGGGACAATTGGAGAAGCGAGCTGTACAAGTACAAGGTGGTCAAGAT
CGAGCCCCTGGGCGTCGCCCCTACCAAGTGCAAGAGAAGAGTCGTGGGCCGTAGAAGGCGGCGGAGAGCT
GTTGGAATTGGCGCCGTGTTCCTGGGCTTTCTGGGAGCCGCTGGATCTACAATGGGCGCTGCCAGCATGA
CCCTGACAGTGCAGGCTAGACAGCTGCTGTCCGGCATTGTCCAGCAGCAGAACAACTGCCTGAGAGCCCC
TGAGTGTCAACAAAGAATGCTGCAGCTGACCGTGTGGGGCATCAAACAGCTGCAGGCCAGAGTTCTGGCC
GTGGAAAGATACCTGGGCGACCAGCAGCTCCTCGGCATCTGGGGATGTTCTGGCAAGCTGATTTGCTGCA
CCGCCGTGCCTTGGAATGCCAGCTGGTCTAACAAGAGCCTGGACCGGATCTGGAACAATATGACCTGGAT
GGAATGGGAGCGCGAGATCGACAACTACACCTCCGAGATCTACACCCTGATCGAGGAAAGCCAGAACCAG
CAAGAGAAAAACGAGCAAGAGCTGCTCGAGCTGGATtaatga

>HV1301405_D1305V1;CON-Schim.6R.DS.SOSIP.664_OPT_D1305V1 (V1 loop from 10.19 isolate)

ATGGGCTCGCTCCAGCCGCTCGCGACGCTGTACCTCCTGGGCATGCTCGTGGCGTCCGTGCTGGCGGCCG
AGAACCTGTGGGTGACGGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCCAACACCACGCTGTTCTGCGC
CAGCGACGCCAAGGCCTACGACACCGAGGTGCACAACGTGTGGGCGACCCACGCCTGCGTGCCGACGGAC
CCCAACCCCCAGGAGATCGTGCTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGG
AGCAGATGCACGAGGACATCATCTCGCTGTGGGACCAGTCCCTGAAGCCGTGCGTGAAGCTGACGCCCCT
GTGCGTGACCCTGAACTGTAGCACCGCCACCGTGAACAACAGGAGCCGTGGACGAGATGAAGAACTGCTCC
TTCAACATCACCACCGAGATCCGCGACAAGAAGCAGAAGGTGTACGCGCTGTTCTACCGGCTGGACGTGG
TGCCGATCGACGACAACAACAACAACTCCAGCAACTACCGCCTGATCAACTGCAACACCAGCGCCTGCAC
CCAGGCCTGCCCGAAGGTGTCCTTCGAGCCCATCCCCATCCACTACTGCGCGCCGGCCGGCTTCGCCATC
CTGAAGTGCAACGACAAGAAGTTCAACGGCACCGGCCCCTGCAAGAACGTGTCCACCGTGCAGTGCACCC
ACGGGATCAAGCCCGTGGTGTCCACGCAGCTGCTGCTGAACGGCTCCCTGGCCGAGGAGGAGATCATCAT
CCGCTCCGAGAACATCACGAACAACGCCAAGACCATCATCGTGCAGCTGAACGAGTCCGTGGAGATCAAC
TGCACCAGGCCCAACAACAACACCCGCAAGTCCATCCGGATCGGCCCTGGCCAGGCGTTCTACGCCACCG
GCGACATCATCGGCGACATCCGCCAGGCGCACTGCAACATCTCGGGCACGAAGTGGAACAAGACCCTGCA
GCAGGTGGCGAAGAAGCTGCGCGAGCACTTCAACAACAAGACCATCATCTTCAAGCCCAGCTCCGGCGGC
GACCTGGAGATCACGACCCACTCCTTCAACTGCCGCGGCGAGTTCTTCTACTGCAACACCTCCGGCCTGT
TCAACTCGACGTGGATCGGGAACGGCACGAAGAACAACAACAACACCAACGACACCATCACCCTGCCCTG

Figure 12 (cont)

CCGCATCAAGCAGATCATCAACATGTGGCAGGGCGTGGGCCAGTGCATGTACGCGCCGCCCATCGAGGGC
AAGATCACCTGCAAGTCCAACATCACCGGCCTGCTCCTGACGCGCGACGGCGGCAACAACAACACCAACG
AGACCGAGATCTTCAGGCCGGGCGGCGGCGACATGCGCGACAACTGGCGCTCGGAGCTGTACAAGTACAA
GGTGGTGAAGATCGAGCCCCTGGGCGTGGCGCCGACGCGCTGCAAGAGACGCGTGGTGGGCCGCAGACGA
AGGAGACGGGCCGTGGGCATCGGCGCGGTGTTCCTGGGCTTCCTGGGAGCAGCTGGTTCGACGATGGGCG
CAGCTTCCATGACCCTGACAGTGCAGGCACGCAACCTGCTCTCCGGCATCGTCCAGCAGCAGTCGAACCT
GCTTCGAGCCCCCGAGGCGCAGCAGCACCTCCTCAAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCA
CGCGTGCTAGCCGTGGAGCGCTACCTCCGCGACCAGCAGCTGCTCGGAATCTGGGGCTGCTCGGGCAAGC
TGATCTGCTGCACCAACGTGCCGTGGAACAGCTCCTGGTCCAACCGCAACCTCTCGGAGATCTGGGACAA
CATGACCTGGCTCCAGTGGGACAAGGAGATCTCGAACTACACCCAGATCATCTACGGCCTGCTGGAGGAG
TCCCAGAACCAGCAGGAGAAGAACGAGCAGGACCTGCTGGCCCTGGACTGATAA

>HV1301580_D230N_H289N_P291S;CH848.3.D1305.10.19_D949V3.DS.SOSIP_D230N_H289N_P291S (glycan holes filled)

ATGGGATCTCTGCAACCTCTGGCCACACTGTACCTGCTGGGAATGCTGGTGGCTTCTGTGCTGGCCGCCG
AGAATCTGTGGGTCACAGTGTACTATGGCGTGCCCGTGTGGAAAGAGGCCAAGACCACACTGTTCTGTGC
CAGCGACGCCAAGGCCTACAAGAAAGAGGTGCACAACGTCTGGGCCACACACGCCTGTGTGCCTACCGAT
CCATCTCCTCAAGAGCTGTTCCTGGAAAACGTGACCGAGAACTTCAACATGTGGAAGAACGACATGGTGG
ACCAGATGCACGAGGACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCTTGCGTGAAGCTGACCCCTCT
GTGCGTGACCCTGATCTGTAGCACCGCCACCGTGAACAACAGAGCCGTGGACGAGATGAAGAACTGCAGC
TTCAACACCACCACCGAGATCCGGGACAAGAAGAAGAAAGAGTACGCCCTGTTCTATCGGAGCGACGTGG
TGCCCCTGGACGAGACAAACAACACCAGCGAGTACCGGCTGATCAACTGCAACACCTCCGCCTGCACTCA
GGCCTGTCCTAAAGTGACCTTCGAGCCCATTCCTATCCACTACTGTGCCCCTGCCGGCTACGCCATCCTG
AAGTGCAACNACGAGACATTCAACGGCACAGGCCCCTGCAGCAATGTGTCCACCGTGCAGTGTACCCACG
GCATCAGACCAGTGGTGTCTACCCAGCTGCTGCTGAATGGAAGCCTGGCCGAGAAAGAAATCGTGATCAG
AAGCGAGAACCTGACCAACAACGCCAAGATCATCATTGTGCATCTGNACACCNCTGTGGAAATCGTGTGC
ACCCGGCCTAACAACAACACCCGGAAGTCTGTGCGGATCGGCCCTGGCCAGACATTCTATGCCACCGGCG
ATATCATCGGCGACATCAAGCAGGCCCACTGCAACATCAGCGAGGAAAAGTGGAACGAGACACTGCAGAA
AGTGGGCATCGAGCTGCAGAAGCACTTCCCCAACAAGACCATCAAGTACAACCAGAGCGCTGGCGGCGAC
ATGGAAATCACCACACACAGCTTCAATTGTGGCGGCGAGTTCTTCTACTGCAATACCAGCAAGCTGTTCA
ACAGCACCTACAACGGCACCTATATCAGCACCAACTCCACCAATAGCACCAGCTACATCACCCTGCAGTG
CCGGATCAAGCAGATCATCAATATGTGGCAAGGCGTCGGCCGGTGTATGTACGCCCCTCCTATCGCCGGC
AACATCACCTGTCGGAGCAATATCACAGGCCTGCTGCTCACCAGAGATGGCGGCATCAACAACGTGTCCA
ACGAGACAGAAACCTTCCGGCCTGCCGGCGGAGACATGAGAGACAATTGGAGAAGCGAGCTGTACAAGTA
CAAGGTGGTCAAGATCGAGCCCCTGGGCGTCGCACCAACACGGTGCAAGAGAAGAGTCGTGGGCCGTCGT
AGAAGGCGGAGAGCCGTTGGAATTGGCGCCGTGTTCCTGGGCTTTCTGGGAGCCGCTGGATCTACAATGG
GCGCTGCCAGCATGACCCTGACAGTGCAGGCTAGAAATCTGCTGAGCGGCATCGTGCAGCAGCAGAGCAA
TCTGCTCAGAGCCCCTGAGGCTCAGCAGCACCTCCTGAAACTGACAGTGTGGGGAATCAAGCAGCTGCAG
GCCAGAGTGCTGGCAGTGGAAAGATACCTGAGGGACCAGCAGCTCCTCGGAATCTGGGGATGTAGCGGCA

Figure 12 (cont)

AGCTGATCTGCTGCACCAACGTGCCCTGGAACTCCAGCTGGTCCAACCGGAATCTGAGCGAGATCTGGGA
TAACATGACCTGGCTGCAGTGGGACAAAGAGATCAGCAACTACACCCAGATCATCTACGGCCTGCTGGAA
GAGAGCCAGAACCAGCAAGAGAAAAACGAGCAGGACCTGCTGGCCCTGGACTGATAA

>HV1301580;CH848.3.D1305.10.19_D949V3.DS.SOSIP (19CV3)

ATGGGATCTCTGCAACCTCTGGCCACACTGTACCTGCTGGGAATGCTGGTGGCTTCTGTGCTGGCCGCCG
AGAATCTGTGGGTCACAGTGTACTATGGCGTGCCCGTGTGGAAAGAGGCCAAGACCACACTGTTCTGTGC
CAGCGACGCCAAGGCCTACAAGAAAGAGGTGCACAACGTCTGGGCCACACACGCCTGTGTGCCTACCGAT
CCATCTCCTCAAGAGCTGTTCCTGGAAAACGTGACCGAGAACTTCAACATGTGGAAGAACGACATGGTGG
ACCAGATGCACGAGGACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCTTGCGTGAAGCTGACCCCTCT
GTGCGTGACCCTGATCTGTAGCACCGCCACCGTGAACAACAGAGCCGTGGACGAGATGAAGAACTGCAGC
TTCAACACCACCACCGAGATCCGGGACAAGAAGAAGAAAGAGTACGCCCTGTTCTATCGGAGCGACGTGG
TGCCCCTGGACGAGACAAACAACACCAGCGAGTACCGGCTGATCAACTGCAACACCTCCGCCTGCACTCA
GGCCTGTCCTAAAGTGACCTTCGAGCCCATTCCTATCCACTACTGTGCCCCTGCCGGCTACGCCATCCTG
AAGTGCAACGACGAGACATTCAACGGCACAGGCCCCTGCAGCAATGTGTCCACCGTGCAGTGTACCCACG
GCATCAGACCAGTGGTGTCTACCCAGCTGCTGCTGAATGGAAGCCTGGCCGAGAAAGAAATCGTGATCAG
AAGCGAGAACCTGACCAACAACGCCAAGATCATCATTGTGCATCTGCACACCCTGTGGAAATCGTGTGC
ACCCGGCCTAACAACAACACCCGGAAGTCTGTGCGGATCGGCCCTGGCCAGACATTCTATGCCACCGGCG
ATATCATCGGCGACATCAAGCAGGCCCACTGCAACATCAGCGAGGAAAAGTGGAACGAGACACTGCAGAA
AGTGGGCATCGAGCTGCAGAAGCACTTCCCCAACAAGACCATCAAGTACAACCAGAGCGCTGGCGGCGAC
ATGGAAATCACCACACACAGCTTCAATTGTGGCGGCGAGTTCTTCTACTGCAATACCAGCAAGCTGTTCA
ACAGCACCTACAACGGCACCTATATCAGCACCAACTCCACCAATAGCACCAGCTACATCACCCTGCAGTG
CCGGATCAAGCAGATCATCAATATGTGGCAAGGCGTCGGCCGGTGTATGTACGCCCCTCCTATCGCCGGC
AACATCACCTGTCGGAGCAATATCACAGGCCTGCTGCTCACCAGAGATGGCGGCATCAACAACGTGTCCA
ACGAGACAGAAACCTTCCGGCCTGCCGGCGGAGACATGAGAGACAATTGGAGAAGCGAGCTGTACAAGTA
CAAGGTGGTCAAGATCGAGCCCCTGGGCGTCGCACCAACACGGTGCAAGAGAAGAGTCGTGGGCCGTCGT
AGAAGGCGGAGAGCCGTTGGAATTGGCGCCGTGTTCCTGGGCTTTCTGGGAGCCGCTGGATCTACAATGG
GCGCTGCCAGCATGACCCTGACAGTGCAGGCTAGAAATCTGCTGAGCGGCATCGTGCAGCAGCAGAGCAA
TCTGCTCAGAGCCCCTGAGGCTCAGCAGCACCTCCTGAAACTGACAGTGTGGGGAATCAAGCAGCTGCAG
GCCAGAGTGCTGGCAGTGGAAAGATACCTGAGGGACCAGCAGCTCCTCGGAATCTGGGGATGTAGCGGCA
AGCTGATCTGCTGCACCAACGTGCCCTGGAACTCCAGCTGGTCCAACCGGAATCTGAGCGAGATCTGGGA
TAACATGACCTGGCTGCAGTGGGACAAAGAGATCAGCAACTACACCCAGATCATCTACGGCCTGCTGGAA
GAGAGCCAGAACCAGCAAGAGAAAAACGAGCAGGACCTGCTGGCCCTGGACTGATAA

>HV1301509;CH0848.3.d1305.10.19gp160

ATGAGAGTGACGGGGATACTGAGGAATTATCCACAATGGTGGATATGGGGCATCTTAGGCTTTTGGATGCTAATG
ACTTGTAATGGGGAAGGAAACTTGTGGGTCACAGTCTACTATGGGGTACCAGTATGGAAAGAAGCAAAAACTACT
CTGTTTTGTGCATCAGATGCCAAAGCGTATAAGAAAGAAGTGCATAATGTCTGGGCGACACATGCCTGTGTACCCA
CAGACCCCAGCCCACAAGAACTGTTTTTTGGAAAATGTAACAGAAAATTTTAACATGTGGAAAAATGATATGGTAG
ATCAGATGCATGAGGATATAATCAGTCTATGGGATCAAAGCCTAAAGCCATGTGTAAAGTTGACCCCACTCTGTGT

Figure 12 (cont)

CACTTTAATTTGTAGTACTGCTACTGTTAACAATAGAGCAGTTGATGAAATGAAAAATTGCTCCTTCAATACAACCA
CAGAAATAAGAGATAAGAAAAAGAAGGAATATGCACTTTTTTATAGATCTGATGTAGTACCACTTGATGAAACAA
ACAATACCAGTGAGTATAGATTAATAAATTGTAATACCTCAGCCGTAACACAAGCCTGTCCAAAGGTCACTTTTGA
ACCAATTCCTATACATTATTGTGCTCCAGCTGGTTATGCGATTCTAAAGTGCAATGATGAGACATTCAATGGAACA
GGACCATGCAGTAATGTCAGCACAGTACAATGTACACATGGAATTAGGCCAGTGGTATCAACCCAACTACTGTTAA
ATGGTAGTCTGGCAGAAAAAGAGATAGTAATTAGATCTGAAAACCTGACAAACAATGCCAAAATAATAATAGTCC
ATCTTCACACCCCTGTAGAAATTGTGTGTACAAGGCCCGGCCATAATACAAGGAAAAGTGTGAGAATAGGCCCAG
GACAAACATTCTATGCAACAGGAGACATAATAGGAGATATAAGACAAGCACATTGTAACATTAATGAAAGTAAAT
GGAATGAAACTTTACAAAAGGTAGGTATAGAATTGCAAAAACACTTCCCTAATAAGACAATAAAGTATAACCAATC
CGCAGGAGGAGACATGGAAATTACAACACATAGCTTTAATTGTGGAGGAGAATTTTTCTATTGCAATACATCAAAG
CTGTTTAATAGTACATACAATGGTACATACATAAGTACAAACAGCACAAACAGTACTTCATACATCACGCTTCAATG
CAGAATAAAACAAATTATAAACATGTGGCAGGGGGTAGGAAGAGCAATGTATGCTCCTCCCATTGCAGGAAACAT
AACATGTAGATCAAATATCACAGGGCTACTATTGACACGTGATGGAGGGATCAATAATGTTAGCAACGAGACAGA
GACATTCAGGCCTGCAGGAGGAGATATGAGGGATAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAGAAG
TTCAGCCATTAGGAATAGCACCAACTGGTGCAAAAAGGAGAGTGGTGGAGAGAGAAAAAAGAGCAGCAGGACT
AGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCAGGAAGCACTATGGGCGCAGCATCAATAACGCTGACGGTACA
GGCCAGACAATTGTTGTCTGGTATAGTGCAACAGCAAAGCAATTTGCTGAGGGCTATAGAGGCGCAACAGCATAT
GTTGCAACTCACGGTCTGGGGCATCAAGCAGCTCCAGGCAAGAGTCCTGGCTCTGGAAAGATACCTAAAGGATCA
ACAGCTCCTAGGGATGTGGGGCTGCTCTGGAAAACTCATCTGCACCACTAATGTGCCTTGGAACACTAGTTGGAGT
AATAAATCTGAAATGGATATTTGGAATAACATGACATGGATGCAGTGGGAAAGAGAAATTAGCAATTACACAGAG
ACAATATACATGTTGCTTGAAGACTCGCAACGCCAGCAGGAGAGAAATGAAAAAGATTTACTAGCATTGGACAGT
TGGAACAGTCTGTGGAATTGGTTTAACATAACAAACTGGCTGTGGTATATAAAAATATTCATAATGATAGTAGGG
GGCTTGATAGGTTTAAGAATAGTTTTTGCTGTGCTATCTATAGTGAATAGAGTCAGGCAGGGATACTCACCTTTGT
CGTTGCAGACCCTTACCCCAAACCCGAGGGAACCCGACAGGCTCAGAGGAATCGAAGAAGAAGGTGGAGAGCAA
GACAGAGACAGATCCATTCGATTAGTGAGCGGATTCTTGCCAATTGTCTGGGACGACCTGCGGAGCCTGTGCCTCT
TCAGTTACCACCGATTGAGAGACTTTCTATTGCTGGCAGCGAGAGTGGTGGAACTTCTGGGACACAGCAGTCTCA
GGGGACTGCAGAGGGGGTGGGAAGTCCTTAAGTATCTGGGAAGTCTTGTGCAGTATTGGGGTCTGGAACTAAAA
AGGAGTGCTATTAGTCTTTTTGATACCCTAGCAATAGCAGTAGCTGAAGGAACAGATAGGATTATAGAATTAATAC
AAGGATTTTGTAGAGCTATCCGCAACATACCTACAAGAATAAGACAAGGCTTTGAAGCATCTTTGCTATAATAATAA

>HV1301503;CH848.3.D1305.10.19ch.DS.SOSIP.664

ATGGGATCTCTGCAGCCTCTGGCCACACTGTACCTGCTGGGAATGCTGGTGGCTTCTGTGCTGGCCGCCG
AGAATCTGTGGGTCACAGTGTACTATGGCGTGCCCGTGTGGAAAGAGGCCAAGACCACACTGTTCTGTGC
CAGCGACGCCAAGGCCTACAAGAAAGAGGTGCACAACGTCTGGGCCACACACGCCTGTGTGCCTACCGAT
CCATCTCCTCAAGAGCTGTTCCTGGAAAACGTGACCGAGAACTTCAACATGTGGAAGAACGACATGGTGG
ACCAGATGCACGAGGACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCTTGCGTGAAGCTGACCCCTCT
GTGCGTGACCCTGATCTGTAGCACCGCCACCGTGAACAACAGAGCCGTGGACGAGATGAAGAACTGCAGC

Figure 12 (cont)

TTCAACACCACCACCGAGATCCGGGACAAGAAGAAGAAAGAGTACGCCCTGTTCTATCGGAGCGACGTGG
TGCCCCTGGACGAGACAAACAACACCAGCGAGTACCGGCTGATCAACTGCAACACCTCCGCCTGCACACA
GGCTTGCCCCAAAGTGACCTTCGAGCCCATTCCTATCCACTACTGTGCCCCTGCCGGCTACGCCATCCTG
AAGTGCAACGACGAGACATTCAACGGCACAGGCCCCTGCAGCAATGTGTCCACCGTGCAGTGTACCCACG
GCATCAGACCAGTGGTGTCTACCCAGCTGCTGCTGAATGGAAGCCTGGCCGAGAAAGAAATCGTGATCAG
AAGCGAGAACCTGACCAACAACGCCAAGATCATCATTGTGCATCTGCACACCCCTGTGGAAATCGTGTGC
ACCAGACCTGGCCACAACACCCGGAAGTCTGTGCGGATTGGACCCGGCCAGACCTTTTATGCCACCGGCG
ATATCATCGGCGACATCAGACAGGCCCACTGTAACATCAACGAGAGCAAGTGGAACGAGACACTGCAGAA
AGTGGGCATCGAGCTGCAGAAGCACTTCCCCAACAAGACCATCAAGTACAACCAGAGCGCTGGCGGCGAC
ATGGAAATCACCACACACAGCTTCAATTGTGGCGGCGAGTTCTTCTACTGCAATACCAGCAAGCTGTTCA
ACAGCACCTACAACGGCACCTATATCAGCACCAACTCCACCAATAGCACCAGCTACATCACCCTGCAGTG
CCGGATCAAGCAGATCATCAATATGTGGCAAGGCGTGGGCAGATGCATGTACGCCCCTCCTATCGCCGGC
AACATCACCTGTCGGAGCAATATCACAGGCCTGCTGCTCACCAGAGATGGCGGCATCAACAACGTGTCCA
ACGAGACAGAAACCTTCCGGCCTGCCGGCGGAGACATGAGAGACAATTGGAGAAGCGAGCTGTACAAGTA
CAAGGTGGTCAAGATCGAGCCCCTGGGCGTCGCACCAACACGGTGCAAGAGAAGAGTCGTGGGACGTAGA
CGAAGGCGGAGAGCCGTTGGAATCGGAGCCGTGTTCCTGGGCTTTCTGGGAGCCGCTGGATCTACAATGG
GCGCTGCCAGCATGACCCTGACAGTGCAGGCTAGAAATCTGCTGAGCGGCATCGTGCAGCAGCAGAGCAA
TCTGCTCAGAGCCCCTGAGGCTCAGCAGCACCTCCTGAAACTGACAGTGTGGGGCATCAAGCAGCTGCAG
GCAAGAGTGCTGGCAGTGGAAAGATACCTGCGGGACCAGCAGCTCCTCGGAATCTGGGGATGTAGCGGCA
AGCTGATCTGCTGCACCAACGTGCCCTGGAACTCCAGCTGGTCCAACCGGAATCTGAGCGAGATCTGGGA
TAACATGACCTGGCTGCAGTGGGACAAAGAGATCAGCAACTACACCCAGATCATCTACGGCCTGCTGGAA
GAGAGCCAGAACCAGCAAGAGAAAAACGAGCAGGACCTGCTGGCCCTGGACTGATAA

>HV1301504;CH848.3.D1305.10.19ch.SOSIPv6

ATGGGATCTCTGCAGCCTCTGGCCACACTGTACCTGCTGGGAATGCTGGTGGCTTCTGTGCTGGCCGCCG
AGAATCTGTGGGTCACAGTGTACTATGGCGTGCCCGTGTGGAAAGAGGCCNNNACCACACTGTTCTGTGC
CAGCGACGCCAAGGCCTACAAGAAANAAGGTGCNCAACGTCTGGGCCACACACNNCTGTGTGCCTACCGAT
CCATCTCCTCAAGAGCTGTTCCTGGAAAACGTGACCGAGAACTTCAACATGTGGAAGAACGACATGGTGG
ACCAGATGCACGAGGACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCTTGCGTGAAGCTGACCCCTCT
GTGCGTGACCCTGATCTGTAGCACCGCCACCGTGAACAACAGAGCCGTGGACGAGATGAAGAACTGCAGC
TTCAACACCACCACCGAGATCCGGGACAAGAAGAAGAAAGAGTACGCCCTGTTCTATCGGAGCGACGTGG
TGCCCCTGGACGAGACAAACAACACCAGCGAGTACCGGCTGATCAACTGCAACACCTCCGCCNNNACACA
GGCTTGCCCCAAAGTGACCTTCGAGCCCATTCCTATCCACTACTGTGCCCCTGCCGGCTACGCCATCCTG
AAGTGCAACGACGAGACATTCAACGGCACAGGCCCCTGCAGCAATGTGTCCACCGTGCAGTGTACCCACG
GCATCAGACCAGTGGTGTCTACCCAGCTGCTGCTGAATGGAAGCCTGGCCGAGAAAGAAATCGTGATCAG
AAGCGAGAACCTGACCAACAACGCCAAGATCATCATTGTGCATCTGCACACCCCTGTGGAAATCGTGTGC
ACCAGACCTGGCCACAACACCCGGAAGTCTGTGCGGATTGGACCCGGCCAGNNNTTTTATGCCACCGGCG
ATATCATCGGCGACATCAGACAGGCCCACTGTAACATCAACGAGAGCAAGTGGAACGAGACACTGCAGAA
AGTGGGCATCGAGCTGCAGAAGCACTTCCCCAACAAGACCATCAAGTACAACCAGAGCGCTGGCGGCGAC

Figure 12 (cont)

ATGGAAATCACCACACACAGCTTCAATTGTGGCGGCGAGTTCTTCTACTGCAATACCAGCAAGCTGTTCA
ACAGCACCTACAACGGCACCTATATCAGCACCAACTCCACCAATAGCACCAGCTACATCACCCTGCAGTG
CCGGATCAAGCAGATCATCAATATGTGGCAAGGCGTGGGCAGAGCCTATGTACGCCCCTCCTATCGCCGGC
AACATCACCTGTCGGAGCAATATCACAGGCCTGCTGCTCACCAGAGATGGCGGCATCAACAACGTGTCCA
ACGAGACAGAAACCTTCCGGCCTGCCGGCGGAGACATGAGAGACAATTGGAGAAGCGAGCTGTACAAGTA
CAAGGTGGTCAAGATCGAGCCCCTGGGCGTCGCACCAACACGGTGCAAGAGAAGAGTCGTGGGACGTAGA
CGAAGGCGGAGAGCCGTTGGAATCGGAGCCGTGTTCCTGGGCTTTCTGGGAGCCGCTGGATCTACAATGG
GCGCTGCCAGCATGACCCTGACAGTGCAGGCTAGAAATCTGCTGAGCGGCATCGTGCAGCAGCAGAGCAA
TTGCCTCAGAGCCCCTGAGTGTCAGCAGCACCTCCTGAAACTGACAGTGTGGGGCATCAAGCAGCTGCAG
GCAAGAGTGCTGGCAGTGGAAAGATACCTGCGGGACCAGCAGCTCCTCGGAATCTGGGGATGTAGCGGCA
AGCTGATCTGCTGCACCAACGTGCCCTGGAACTCCAGCTGGTCCAACCGGAATCTGAGCGAGATCTGGGA
TAACATGACCTGGCTGCAGTGGGACAAAGAGATCAGCAACTACACCCAGATCATCTACGGCCTGCTGGAA
GAGAGCCAGAACCAGCAAGAGAAAAACGAGCAGGACCTGCTGGCCCTGGACTGATAA

Figure 12 (cont)

Translate results

>HV1301580_D230N_H289N_P291S;CH848.3.D1305.10.19_D949V3.DS.SOSIP_D230N_H289N_P291S

MGSLQPLATLYLLGMLVASVLAAENLWVTVYYGVPVWKEAKTTLFCASDAKAYKKEVHNV
WATHACVPTDPSPQELFLENVTENFNMWKNDMVDQMHEDIISLWDQSLKPCVKLTPLCVT
LICSTATVNNRAVDEMKNCSFNTTTEIRDKKKKEYALFYRSDVVPLDETNNTSEYRLINC
NTSACTQACPKVTFEPIPIHYCAPAGYAILKCNNETFNGTGPCSNVSTVQCTHGIRPVVS
TQLLLNGSLAEKEIVIRSENLTNNAKIIIVHLNTSVEIVCTRPNNNTRKSVRIGPGQTFY
ATGDIIGDIKQAHCNISEEKWNETLQKVGIELQKHFPNKTIKYNQSAGGDMEITTHSFNC
GGEFFYCNTSKLFNSTYNGTYISTNSTNSTSYITLQCRIKQIINMWQGVGRCMYAPPIAG
NITCRSNITGLLLTRDGGINNVSNETETFRPAGGDMRDNWRSELYKYKVVKIEPLGVAPT
RCKRRVVGRRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLR
APEAQQHLLKLTVWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNR
NLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALD**

>HV1301502_D1305V1;JRFL_SOSIPv6_V1_PNGS_D1305V1

MGSLQPLATLYLLGMLVASCLGVEKLWVTVYYGVPVWKEACTTLFCASDAKAYDTKVRNV
WATHCCVPTDPNPQEVVLENVTEHFNMWKNNMVEQMQEDIISLWDQSLKPCVKLTPLCVT
LNCSTATVNNRAVDEMKNCSFNITTSIRDKVQKEYALFYKLDVVPIDNNNTSYRLISCDT
SVITQACPKISFEPIPIHYCAPAGFAILKCNDKTFNGKGPCKNVSTVQCTHGIRPVVSTQ
LLLNGSLAEEEVVIRSDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGPGRWFYTT
GEIIGDIRQAHCNISRAKWNDTLKQIVIKLREQFENKTIVFNHSSGGDPEIVMHSFNCGG
EFFYCNSTQLFNSTWNNNTEGSNNTEGNTITLPCRIKQIINMWQEIGKAMYAPPIRGQIR
CSSNITGLLLTRDGGINENGTEIFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTKCKRR
VVGRRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARQLLSGIVQQQNNCLRAPECQ
QRMLQLTVWGIKQLQARVLAVERYLGDQQLLGIWGCSGKLICCTAVPWNASWSNKSLDRI
WNNMTWMEWEREIDNYTSEIYTLIEESQNQQEKNEQELLELD**

>HV1301405_D1305V1;CON-Schim.6R.DS.SOSIP.664_OPT_D1305V1

MGSLQPLATLYLLGMLVASVLAAENLWVTVYYGVPVWKEANTTLFCASDAKAYDTEVHNV
WATHACVPTDPNPQEIVLENVTENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVT
LNCSTATVNNRAVDEMKNCSFNITTEIRDKKQKVYALFYRLDVVPIDDNNNNSSNYRLIN
CNTSACTQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGTGPCKNVSTVQCTHGIKPVV
STQLLLNGSLAEEEIIIRSENITNNAKTIIVQLNESVEINCTRPNNNTRKSIRIGPGQAF
YATGDIIGDIRQAHCNISGTKWNKTLQQVAKKLREHFNNKTIIFKPSSGGDLEITTHSFN
CRGEFFYCNTSGLFNSTWIGNGTKNNNNTNDTITLPCRIKQIINMWQGVGQCMYAPPIEG
KITCKSNITGLLLTRDGGNNNTNETEIFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTR
CKRRVVGRRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRA
PEAQQHLLKLTVWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRN
LSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALD**

Figure 13

>HV1301580_D230N_H289N_P291S;CH848.3.D1305.10.19_D949V3.DS.SOSIP_D230N_H289N_P291S

MGSLQPLATLYLLGMLVASVLAAENLWVTVYYGVPVWKEAKTTLFCASDAKAYKKEVHNV
WATHACVPTDPSPQELFLENVTENFNMWKNDMVDQMHEDIISLWDQSLKPCVKLTPLCVT
LICSTATVNNRAVDEMKNCSFNTTTEIRDKKKKEYALFYRSDVVPLDETNNTSEYRLINC
NTSACTQACPKVTFEPIPIHYCAPAGYAILKCNNETFNGTGPCSNVSTVQCTHGIRPVVS
TQLLLNGSLAEKEIVIRSENLTNNAKIIIVHLNTSVEIVCTRPNNNTRKSVRIGPGQTFY
ATGDIIGDIKQAHCNISEEKWNETLQKVGIELQKHFPNKTIKYNQSAGGDMEITTHSFNC
GGEFFYCNTSKLFNSTYNGTYISTNSTNSTSYITLQCRIKQIINMWQGVGRCMYAPPIAG
NITCRSNITGLLLTRDGGINNVSNETETFRPAGGDMRDNWRSELYKYKVVKIEPLGVAPT
RCKRRVVGRRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLR
APEAQQHLLKLTVWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNR
NLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALD**

>HV1301580;CH848.3.D1305.10.19_D949V3.DS.SOSIP (19CV3)

MGSLQPLATLYLLGMLVASVLAAENLWVTVYYGVPVWKEAKTTLFCASDAKAYKKEVHNV
WATHACVPTDPSPQELFLENVTENFNMWKNDMVDQMHEDIISLWDQSLKPCVKLTPLCVT
LICSTATVNNRAVDEMKNCSFNTTTEIRDKKKKEYALFYRSDVVPLDETNNTSEYRLINC
NTSACTQACPKVTFEPIPIHYCAPAGYAILKCNDETFNGTGPCSNVSTVQCTHGIRPVVS
TQLLLNGSLAEKEIVIRSENLTNNAKIIIVHLHTPVEIVCTRPNNNTRKSVRIGPGQTFY
ATGDIIGDIKQAHCNISEEKWNETLQKVGIELQKHFPNKTIKYNQSAGGDMEITTHSFNC
GGEFFYCNTSKLFNSTYNGTYISTNSTNSTSYITLQCRIKQIINMWQGVGRCMYAPPIAG
NITCRSNITGLLLTRDGGINNVSNETETFRPAGGDMRDNWRSELYKYKVVKIEPLGVAPT
RCKRRVVGRRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLR
APEAQQHLLKLTVWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNR
NLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALD**

>HV1301509;CH0848.3.d1305.10.19gp160 (natural env)

MRVTGILRNYPQWWIWGILGFWMLMTCNGEGNLWVTVYYGVPVWKEAKTTLFCASDAKAY
KKEVHNVWATHACVPTDPSPQELFLENVTENFNMWKNDMVDQMHEDIISLWDQSLKPCVK
LTPLCVTLICSTATVNNRAVDEMKNCSFNTTTEIRDKKKKEYALFYRSDVVPLDETNNTS
EYRLINCNTSAVTQACPKVTFEPIPIHYCAPAGYAILKCNDETFNGTGPCSNVSTVQCTH
GIRPVVSTQLLLNGSLAEKEIVIRSENLTNNAKIIIVHLHTPVEIVCTRPGHNTRKSVRI
GPGQTFYATGDIIGDIRQAHCNINESKWNETLQKVGIELQKHFPNKTIKYNQSAGGDMEI
TTHSFNCGGEFFYCNTSKLFNSTYNGTYISTNSTNSTSYITLQCRIKQIINMWQGVGRAM
YAPPIAGNITCRSNITGLLLTRDGGINNVSNETETFRPAGGDMRDNWRSELYKYKVVEVQ
PLGIAPTGAKRRVVEREKRAAGLGALFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQQ
SNLLRAIEAQQHMLQLTVWGIKQLQARVLALERYLKDQQLLGMWGCSGKLICTTNVPWNT
SWSNKSEMDIWNNMTWMQWEREISNYTETIYMLLEDSQRQQERNEKDLLALDSWNSLWNW
FNITNWLWYIKIFIMIVGGLIGLRIVFAVLSIVNRVRQGYSPLSLQTLTPNPREPDRLRG
IEEEGGEQDRDRSIRLVSGFLPIVWDDLRSLCLFSYHRLRDFLLLAARVVELLGHSSLRG
LQRGWEVLKYLGSLVQYWGLELKRSAISLFDTLAIAVAEGTDRIIELIQGFCRAIRNIPT
RIRQGFEASLL**

Figure 13 (cont)

>HV1301503;CH848.3.D1305.10.19ch.DS.SOSIP.664 ("natural" env SOSIP)

MGSLQPLATLYLLGMLVASVLAAENLWVTVYYGVPVWKEAKTTLFCASDAKAYKKEVHNV
WATHACVPTDPSPQELFLENVTENFNMWKNDMVDQMHEDIISLWDQSLKPCVKLTPLCVT
LICSTATVNNRAVDEMKNCSFNTTTEIRDKKKKEYALFYRSDVVPLDETNNTSEYRLINC
NTSACTQACPKVTFEPIPIHYCAPAGYAILKCNDETFNGTGPCSNVSTVQCTHGIRPVVS
TQLLLNGSLAEKEIVIRSENLTNNAKIIIVHLHTPVEIVCTRPGHNTRKSVRIGPGQTFY
ATGDIIGDIRQAHCNINESKWNETLQKVGIELQKHFPNKTIKYNQSAGGDMEITTHSFNC
GGEFFYCNTSKLFNSTYNGTYISTNSTNSTSYITLQCRIKQIINMWQGVGRCMYAPPIAG
NITCRSNITGLLLTRDGGINNVSNETETFRPAGGDMRDNWRSELYKYKVVKIEPLGVAPT
RCKRRVVGRRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLR
APEAQQHLLKLTVWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNR
NLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALD**

>HV1301504;CH848.3.D1305.10.19ch.SOSIPv6 ("natural" env SOSIP)

MGSLQPLATLYLLGMLVASVLAAENLWVTVYYGVPVWKEACTTLFCASDAKAYKKKVRNV
WATHCCVPTDPSPQELFLENVTENFNMWKNDMVDQMHEDIISLWDQSLKPCVKLTPLCVT
LICSTATVNNRAVDEMKNCSFNTTTEIRDKKKKEYALFYRSDVVPLDETNNTSEYRLINC
NTSAVTQACPKVTFEPIPIHYCAPAGYAILKCNDETFNGTGPCSNVSTVQCTHGIRPVVS
TQLLLNGSLAEKEIVIRSENLTNNAKIIIVHLHTPVEIVCTRPGHNTRKSVRIGPGQWFY
ATGDIIGDIRQAHCNINESKWNETLQKVGIELQKHFPNKTIKYNQSAGGDMEITTHSFNC
GGEFFYCNTSKLFNSTYNGTYISTNSTNSTSYITLQCRIKQIINMWQGVGRAMYAPPIAG
NITCRSNITGLLLTRDGGINNVSNETETFRPAGGDMRDNWRSELYKYKVVKIEPLGVAPT
RCKRRVVGRRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNCLR
APECQQHLLKLTVWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNR
NLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALD**

Figure 13 (cont)

>HV1301521(originally HV1301405_N130D_N135K_N138S_N141S)CON-Schim.6R.DS.SOSIP.664_OPT_N130D_N135K_N138S_N141S

MGSLQPLATLYLLGMLVASVLAAENLWVTVYYGVPVWKEANTTLFCASDAKAYDTEVHNVWATHACVPTDPNPQEI
VLENVTENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLDCTNVKVTSTTSNTEEKGEIKNCSFNITTEIRDK
KQKVYALFYRLDVVPIDDNNNNSSNYRLINCNTSACTQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGTGPCKNVSTV
QCTHGIKPVVSTQLLLNGSLAEEEIIIRSENITNNAKTIIVQLNESVEINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAH
CNISGTKWNKTLQQVAKKLREHFNNKTIIFKPSSGGDLEITTHSFNCRGEFFYCNTSGLFNSTWIGNGTKNNNNTNDTIT
LPCRIKQIINMWQGVGQCMYAPPIEGKITCKSNITGLLLTRDGGNNNTNETEIFRPGGGDMRDNWRSELYKYKVVKIEP
LGVAPTRCKRRVVGRRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQQHLLKLT
VWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLL
EESQNQQEKNEQDLLALD**

>HV1301405_N138S_N141S;CON-Schim.6R.DS.SOSIP.664_OPT_N138S_N141S
MGSLQPLATLYLLGMLVASVLAAENLWVTVYYGVPVWKEANTTLFCASDAKAYDTEVHNV
WATHACVPTDPNPQEIVLENVTENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVT
LNCTNVNVTSTTSNTEEKGEIKNCSFNITTEIRDKKQKVYALFYRLDVVPIDDNNNNSSN
YRLINCNTSACTQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGTGPCKNVSTVQCTHG
IKPVVSTQLLLNGSLAEEEIIIRSENITNNAKTIIVQLNESVEINCTRPNNNTRKSIRIG
PGQAFYATGDIIGDIRQAHCNISGTKWNKTLQQVAKKLREHFNNKTIIFKPSSGGDLEIT
THSFNCRGEFFYCNTSGLFNSTWIGNGTKNNNNTNDTITLPCRIKQIINMWQGVGQCMYA
PPIEGKITCKSNITGLLLTRDGGNNNTNETEIFRPGGGDMRDNWRSELYKYKVVKIEPLG
VAPTRCKRRVVGRRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQS
NLLRAPEAQQHLLKLTVWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNVPWNSS
WSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALD**

Figure 14

>HV1301580_C_SORTA;CH848.3.D1305.10.19_D949V3.DS.SOSIP_C_SORTA

MPMGSLQPLATLYLLGMLVASVLAAENLWVTVYYGVPVWKEAKTTLFCASDAKAYKKEVHNVWATHACVP
TDPSPQELFLENVTENFNMWKNDMVDQMHEDIISLWDQSLKPCVKLTPLCVTLICSTATVNNRAVDEMKN
CSFNTTTEIRDKKKKEYALFYRSDVVPLDETNNTSEYRLINCNTSACTQACPKVTFEPIPIHYCAPAGYA
ILKCNDETFNGTGPCSNVSTVQCTHGIRPVVSTQLLLNGSLAEKEIVIRSENLTNNAKIIIVHLHTPVEI
VCTRPNNNTRKSVRIGPGQTFYATGDIIGDIKQAHCNISEEKWNETLQKVGIELQKHFPNKTIKYNQSAG
GDMEITTHSFNCGGEFFYCNTSKLFNSTYNGTYISTNSTNSTSYITLQCRIKQIINMWQGVGRCMYAPPI
AGNITCRSNITGLLLTRDGGINNVSNETETFRPAGGDMRDNWRSELYKYKVVKIEPLGVAPTRCKRRVVG
RRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQQHLLKLTVWGIKQ
LQARVLAVERYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGL
LEESQNQQEKNEQDLLALDLPSTGG**

>HV1301580_C_SORTA;CH848.3.D1305.10.19_D949V3.DS.SOSIP_C_SORTA

GTCGACGCCACCATGCCCATGGGATCTCTGCAACCTCTGGCCACACTGTACCTGCTGGGAATGCTGGTGGCTTCTGTGCTGGCC
GCCGAGAATCTGTGGGTCACAGTGTACTATGGCGTGCCCGTGTGGAAAGAGGCCAAGACCACACTGTTCTGTGCCAGCGACG
CCAAGGCCTACAAGAAAGAGGTGCACAACGTCTGGGCCACACACGCCTGTGTGCCTACCGATCCATCTCCTCAAGAGCTGTTC
CTGGAAAACGTGACCGAGAACTTCAACATGTGGAAGAACGACATGGTGGACCAGATGCACGAGGACATCATCAGCCTGTGGG
ACCAGAGCCTGAAGCCTTGCGTGAAGCTGACCCCTCTGTGCGTGACCCTGATCTGTAGCACCGCCACCGTGAACAACAGAGCC
GTGGACGAGATGAAGAACTGCAGCTTCAACACCACCACCGAGATCCGGGACAAGAAGAAGAAAGAGTACGCCCTGTTCTATC
GGAGCGACGTGGTGCCCCTGGACGAGACAAACAACACCAGCGAGTACCGGCTGATCAACTGCAACACCTCCGCCTGCACTCA
GGCCTGTCCTAAAGTGACCTTCGAGCCCATTCCTATCCACTACTGTGCCCCTGCCGGCTACGCCATCCTGAAGTGCAACGACGA
GACATTCAACGGCACAGGCCCCTGCAGCAATGTGTCCACCGTGCAGTGTACCCACGGCATCAGACCAGTGGTGTCTACCCAGC
TGCTGCTGAATGGAAGCCTGGCCGAGAAAGAAATCGTGATCAGAAGCGAGAACCTGACCAACAACGCCAAGATCATCATTGT
GCATCTGCACACCCCTGTGGAAATCGTGTGCACCCGGCCTAACAACAACACCCGGAAGTCTGTGCGGATCGGCCCTGGCCAGA
CATTCTATGCCACCGGCGATATCATCGGCGACATCAAGCAGGCCCACTGCAACATCAGCGAGGAAAAGTGGAACGAGACACT
GCAGAAAGTGGGCATCGAGCTGCAGAAGCACTTCCCCAACAAGACCATCAAGTACAACCAGAGCGCTGGCGGCGACATGGAA
ATCACCACACACAGCTTCAATTGTGGCGGCGAGTTCTTCTACTGCAATACCAGCAAGCTGTTCAACAGCACCTACAACGGCACC
TATATCAGCACCAACTCCACCAATAGCACCAGCTACATCACCCTGCAGTGCCGGATCAAGCAGATCATCAATATGTGGCAAGGC
GTCGGCCGGTGTATGTACGCCCCTCCTATCGCCGGCAACATCACCTGTCGGAGCAATATCACAGGCCTGCTGCTCACCAGAGA
TGGCGGCATCAACAACGTGTCCAACGAGACAGAAACCTTCCGGCCTGCCGGCGGAGACATGAGAGACAATTGGAGAAGCGA
GCTGTACAAGTACAAGGTGGTCAAGATCGAGCCCCTGGGCGTCGCACCAACACGGTGCAAGAGAAGAGTCGTGGGCCGTCGT
AGAAGGCGGAGAGCCGTTGGAATTGGCGCCGTGTTCCTGGGCTTTCTGGGAGCCGCTGGATCTACAATGGGCGCTGCCAGCA
TGACCCTGACAGTGCAGGCTAGAAATCTGCTGAGCGGCATCGTGCAGCAGCAGAGCAATCTGCTCAGAGCCCCTGAGGCTCA
GCAGCACCTCCTGAAACTGACAGTGTGGGGAATCAAGCAGCTGCAGGCCAGAGTGCTGGCAGTGGAAAGATACCTGAGGGA
CCAGCAGCTCCTCGGAATCTGGGGATGTAGCGGCAAGCTGATCTGCTGCACCAACGTGCCCTGGAACTCCAGCTGGTCCAACC
GGAATCTGAGCGAGATCTGGGATAACATGACCTGGCTGCAGTGGGACAAAGAGATCAGCAACTACACCCAGATCATCTACGG
CCTGCTGGAAGAGAGCCAGAACCAGCAAGAGAAAAACGAGCAGGACCTGCTGGCCCTGGACCTGCCTAGCACCGGAGGATG
ATAAGGATCC

Figure 15

>CON-S gp160

MRVRGIQRNCQHLWRWGTLILGMLMICSAAENLWVTVYYGVPVWKEANTTLFCASDAKAYDTEVHNVWAT
HACVPTDPNPQEIVLENVTENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTNVNVTNTTN
NTEEKGEIKNCSFNITTEIRDKKQKVYALFYRLDVVPIDDNNNNSSNYRLINCNTSAITQACPKVSFEPI
PIHYCAPAGFAILKCNDKKFNGTGPCKNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENITNNAKT
IIVQLNESVEINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNISGTKWNKTLQQVAKKLREHFN
NKTIIFKPSSGGDLEITTHSFNCRGEFFYCNTSGLFNSTWIGNGTKNNNNTNDTITLPCRIKQIINMWQG
VGQAMYAPPIEGKITCKSNITGLLLTRDGGNNNTNETEIFRPGGGDMRDNWRSELYKYKVVKIEPLGVAP
TKAKRRVVEREKRAVGIGAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLRAIEAQQHLLQL
TVWGIKQLQARVLAVERYLKDQQLLGIWGCSGKLICTTTVPWNSSWSNKSQDEIWDNMTWMEWEREINNY
TDIIYSLIEESQNQQEKNEQELLALDKWASLWNWFDITNWLWYIKIFIMIVGGLIGLRIVFAVLSIVNRV
RQGYSPLSFQTLIPNPRGPDRPEGIEEEGGEQDRDRSIRLVNGFLALAWDDLRSLCLFSYHRLRDFILIA
ARTVELLGRKGLRRGWEALKYLWNLLQYWGQELKNSAISLLDTTAIAVAEGTDRVIEVVQRACRAILNIP
RRIRQGLERALL**

Figure 16

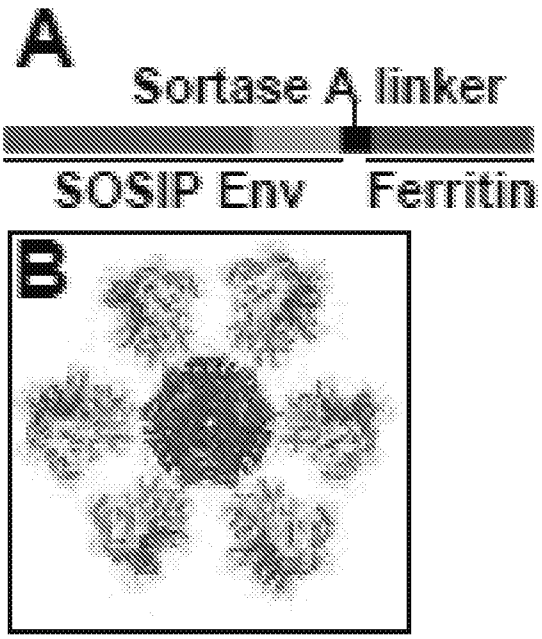


Figure 17 A and B

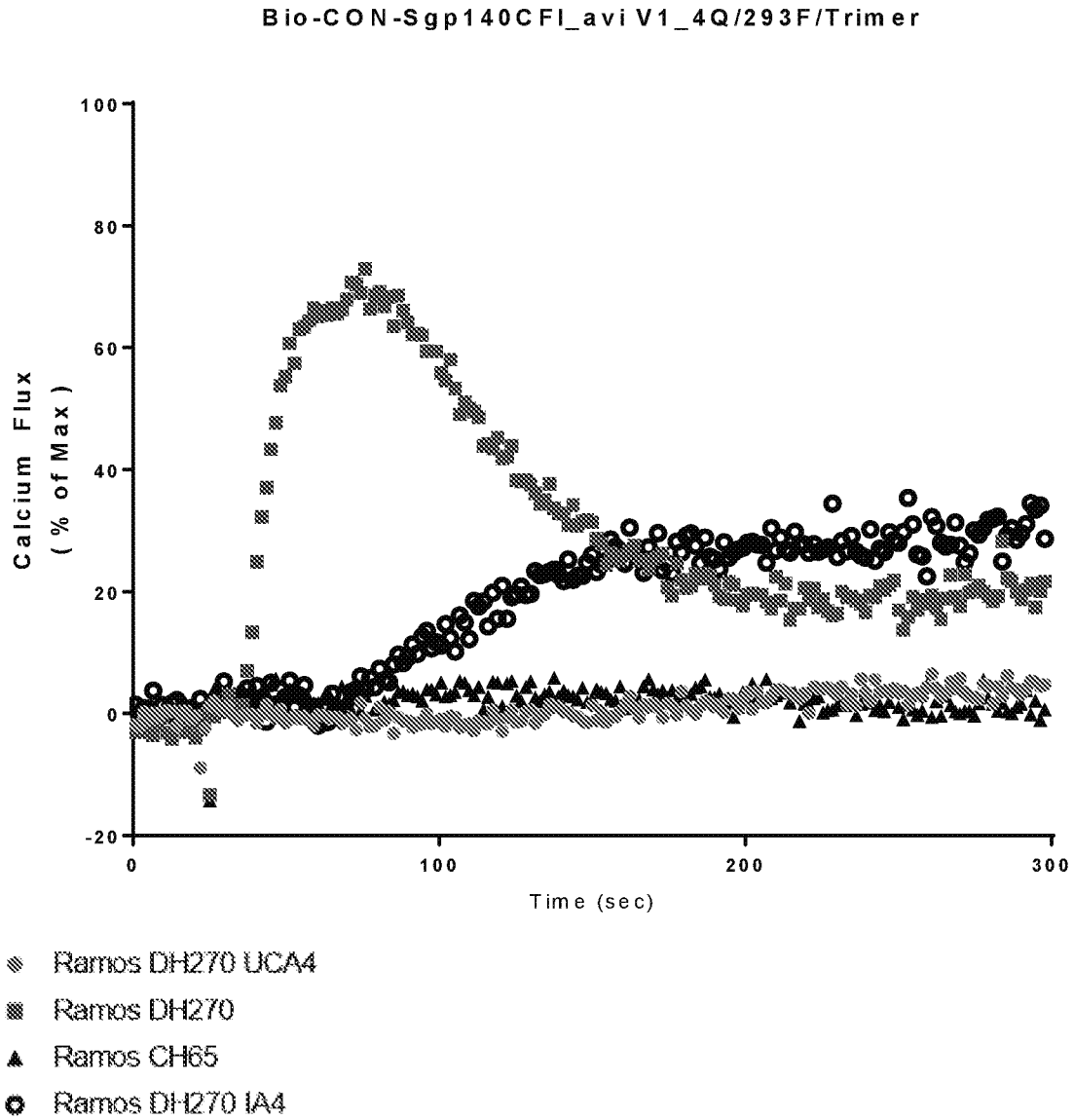

Bio-CON-Sgp140CFI_avi V1_4Q/293F/Trimer
Calcium Flux
(% of Max)
100
80
60
40
20
0
-20
0
100
200
300
Time (sec)
Ramos DH270 UCA4
Ramos DH270
Ramos CH65
Ramos DH270 IA4

Figure 18A

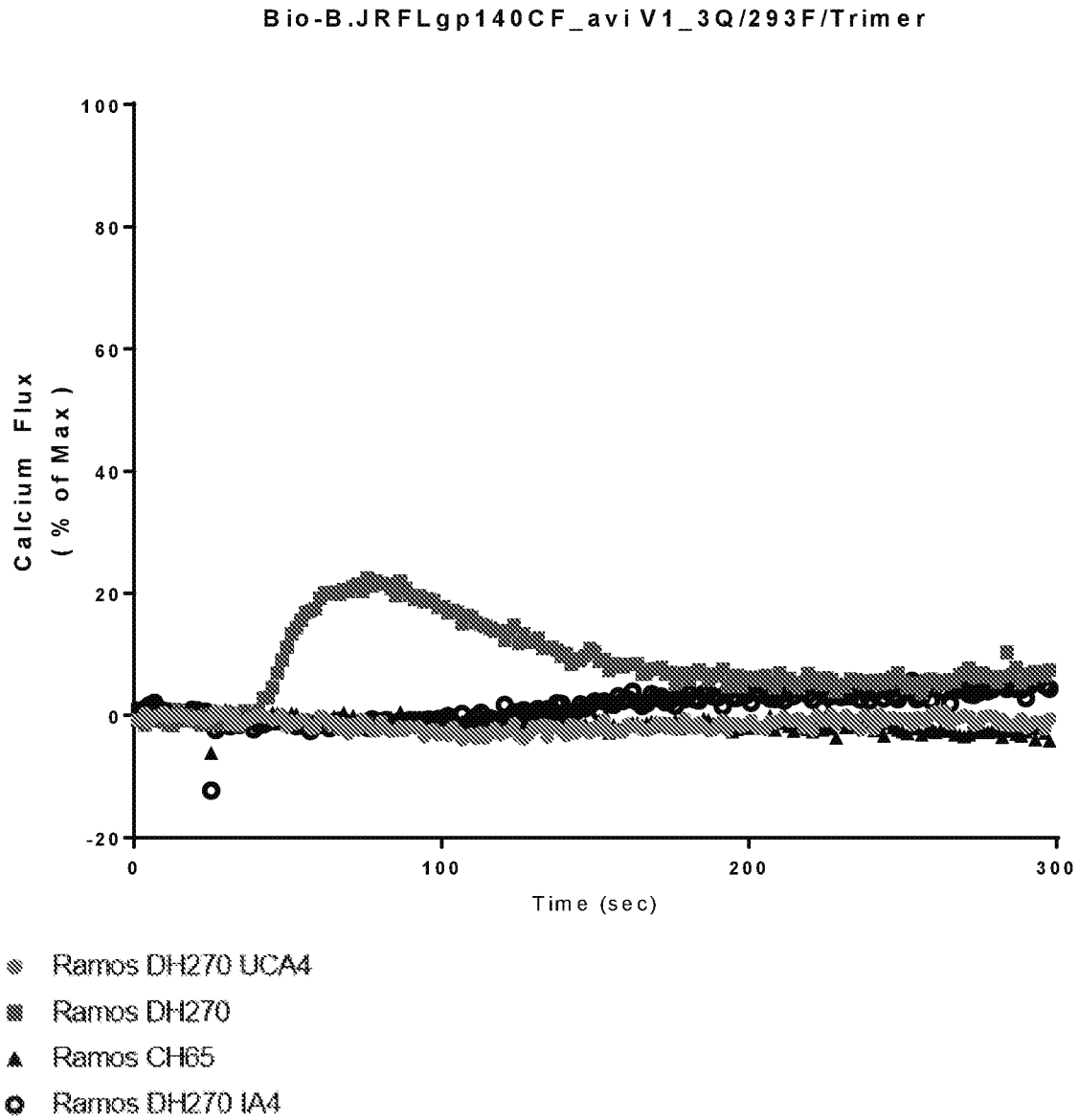

Bio-B.JRFLgp140CF_avi V1_3Q/293F/Trimer
Calcium Flux
(% of Max)
100
80
60
40
20
0
-20
0
100
200
300
Time (sec)
Ramos DH270 UCA4
Ramos DH270
Ramos CH65
Ramos DH270 IA4

Figure 18B

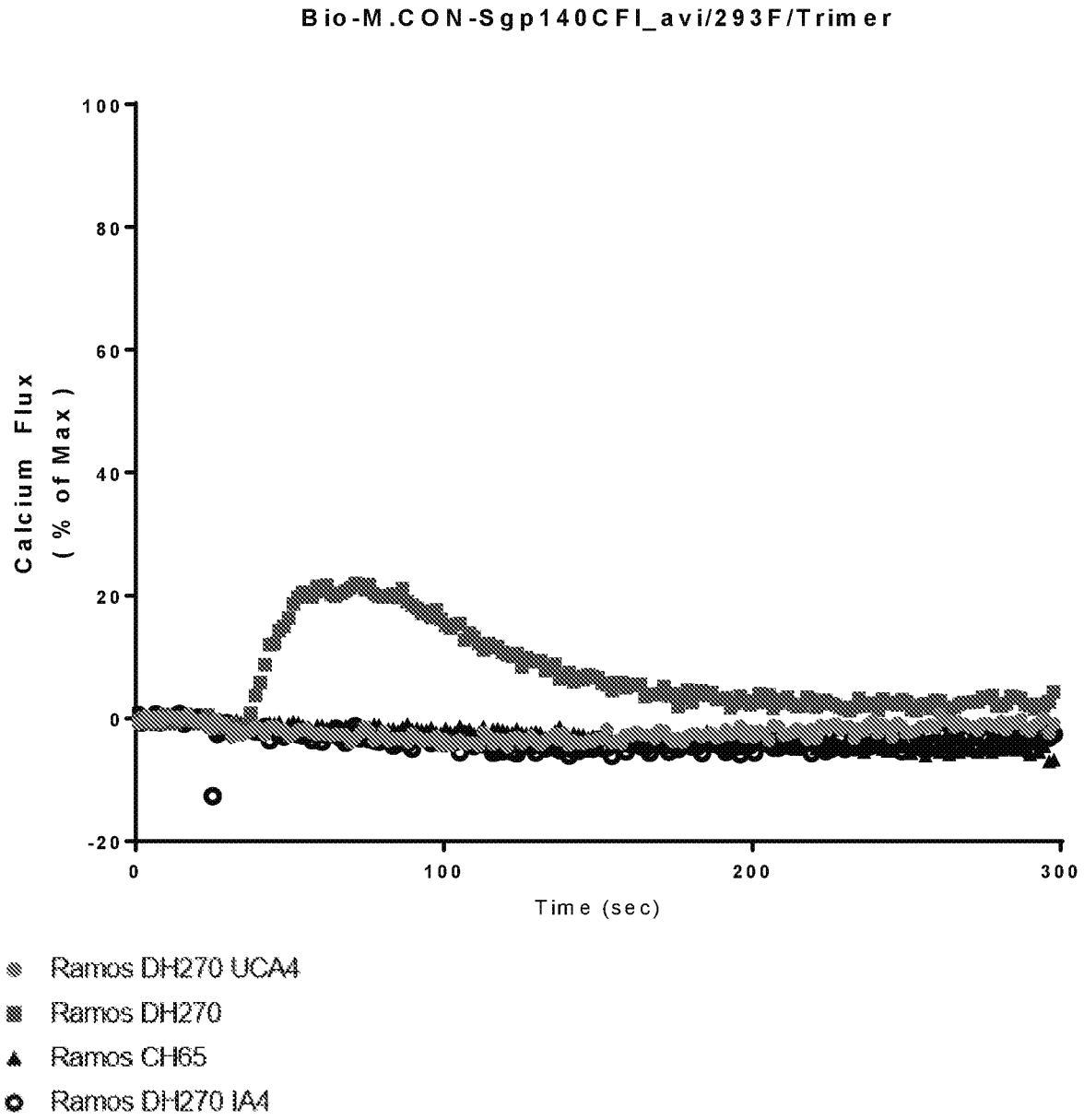
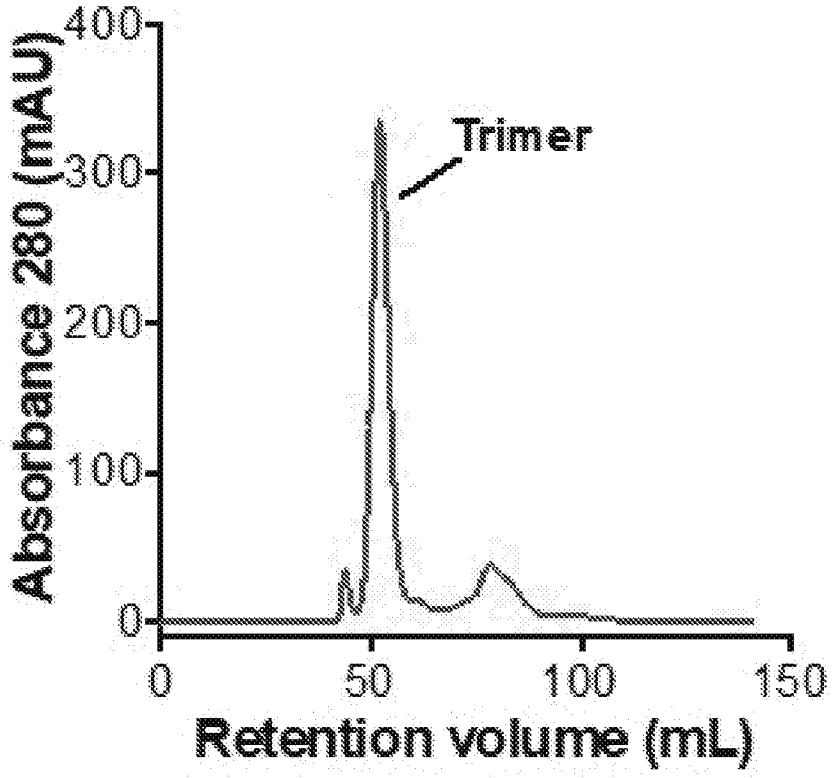

Bio-M.CON-Sgp140CFI_avi/293F/Trimer
Calcium Flux
(% of Max)
100
80
60
40
20
0
-20
0
100
200
300
Time (sec)
Ramos DH270 UCA4
Ramos DH270
Ramos CH65
Ramos DH270 IA4

Figure 18C


Absorbance 280 (mAU)
400
300
200
100
0
Trimer
0
50
100
150
Retention volume (mL)

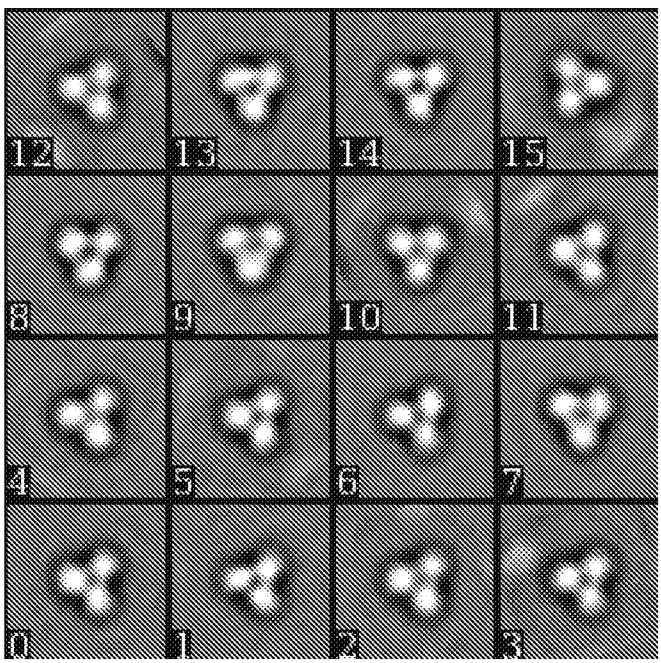
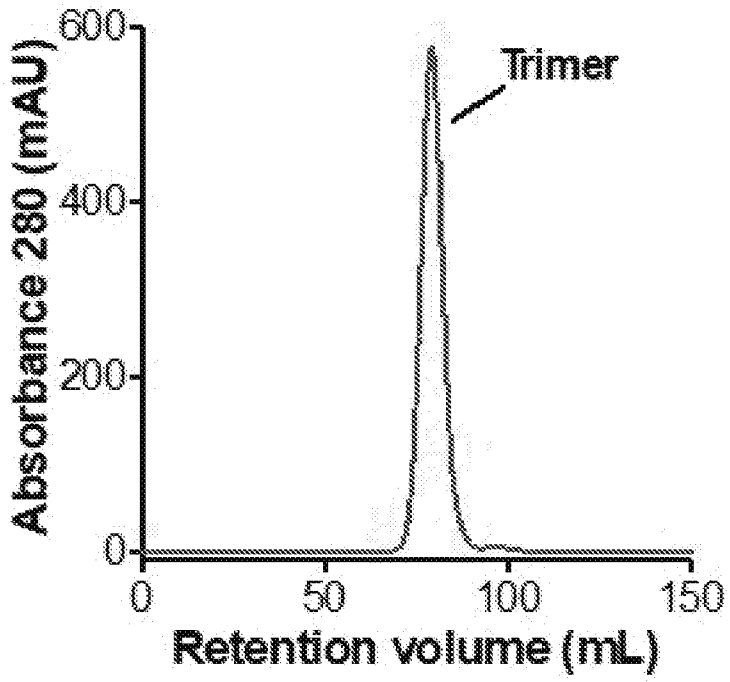

12
13
14
15
8
9
10
11
4
5
6
7
0
1
2
3

Figure 19A


600
400
200
0
Absorbance 280 (mAU)
Trimer
0
50
100
150
Retention volume (mL)

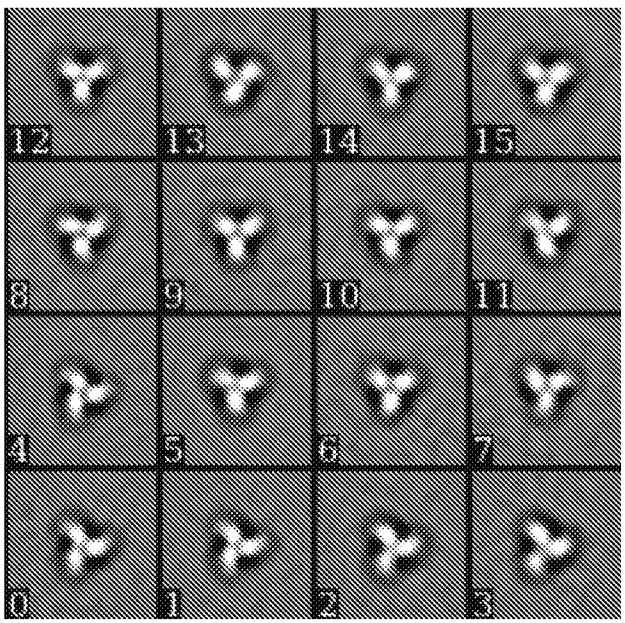
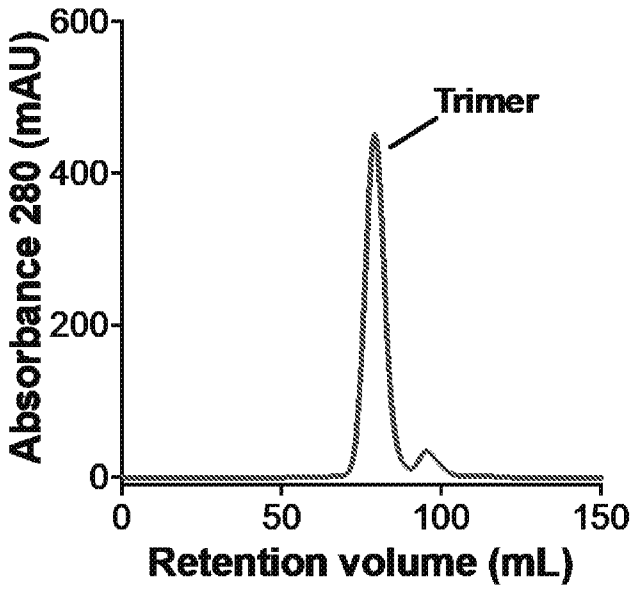

12
13
14
15
8
9
10
11
4
5
6
7
0
1
2
3

Figure 19B


600
400
200
0
Absorbance 280 (mAU)
Trimer
0
50
100
150
Retention volume (mL)

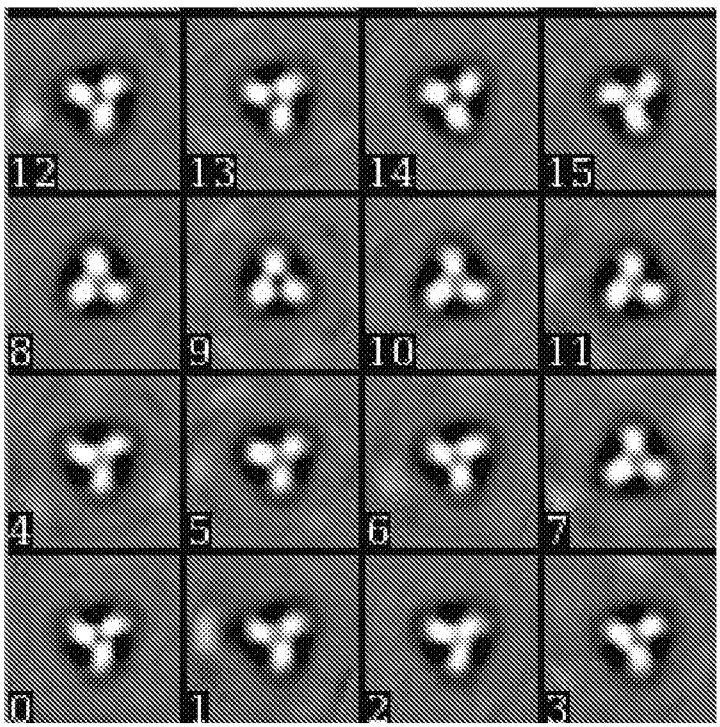

12
13
14
15
8
9
10
11
4
5
6
7
0
1
2
3

Figure 19C

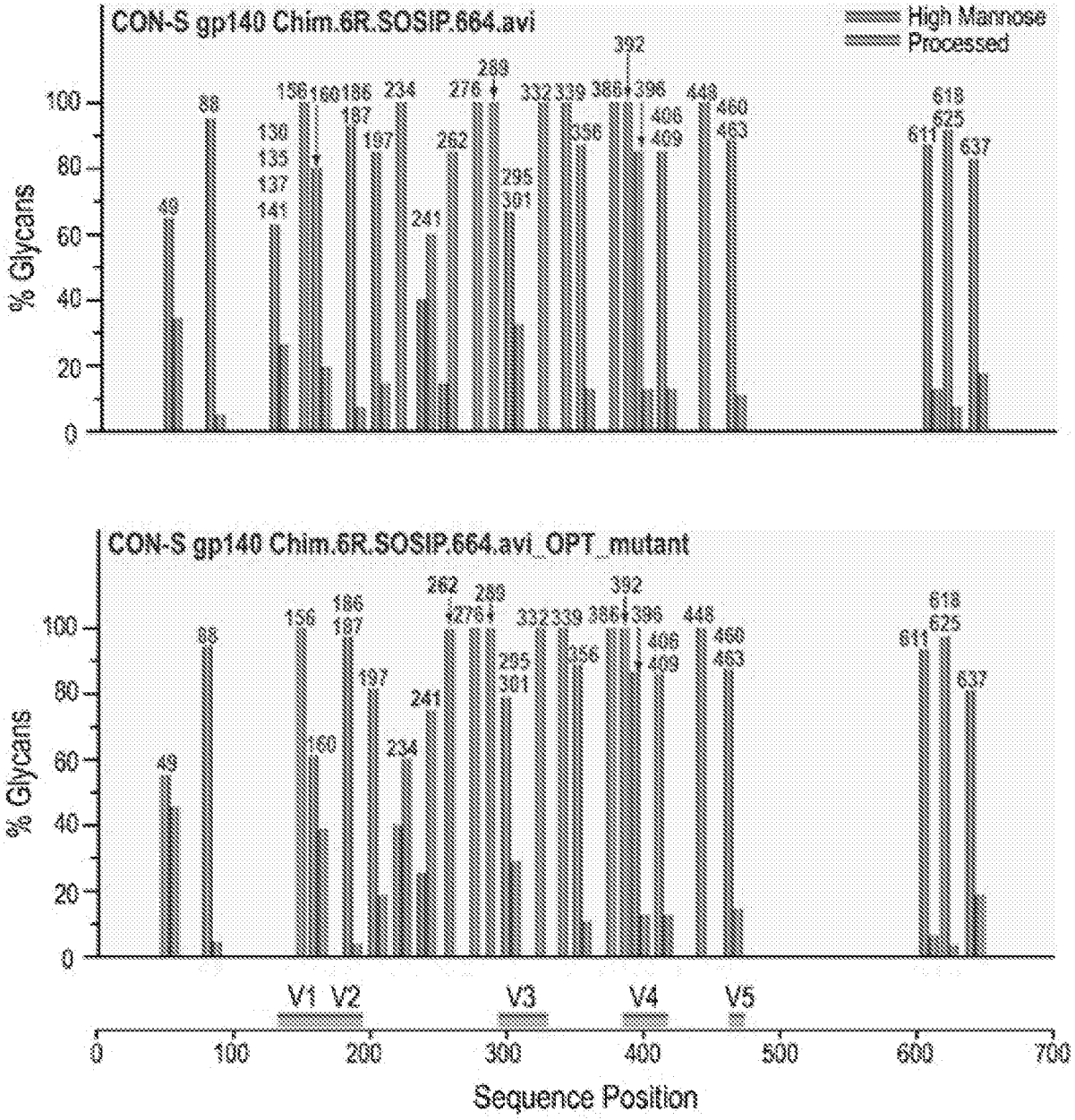

CON-S gp140 Chim.6R.SOSIP.664.avi
High Mannose
Processed
% Glycans
CON-S gp140 Chim.6R.SOSIP.664.avi_OPT_mutant
V1 V2
V3
V4
V5
Sequence Position

Figure 20

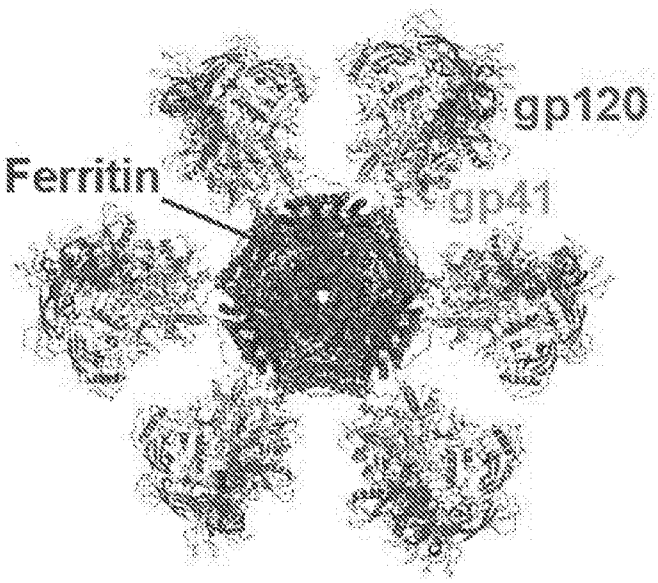


Figure 21A

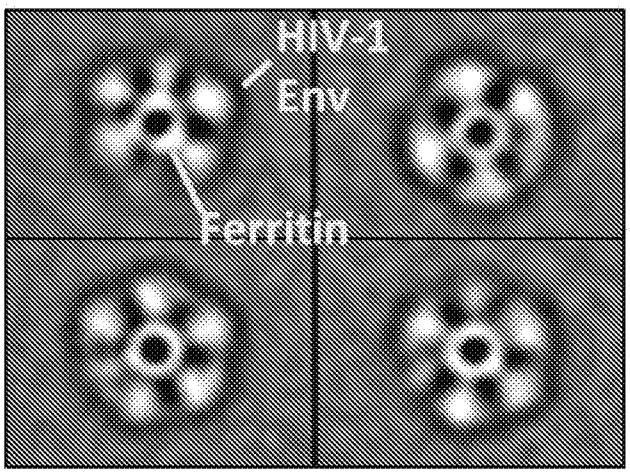


Figure 21B

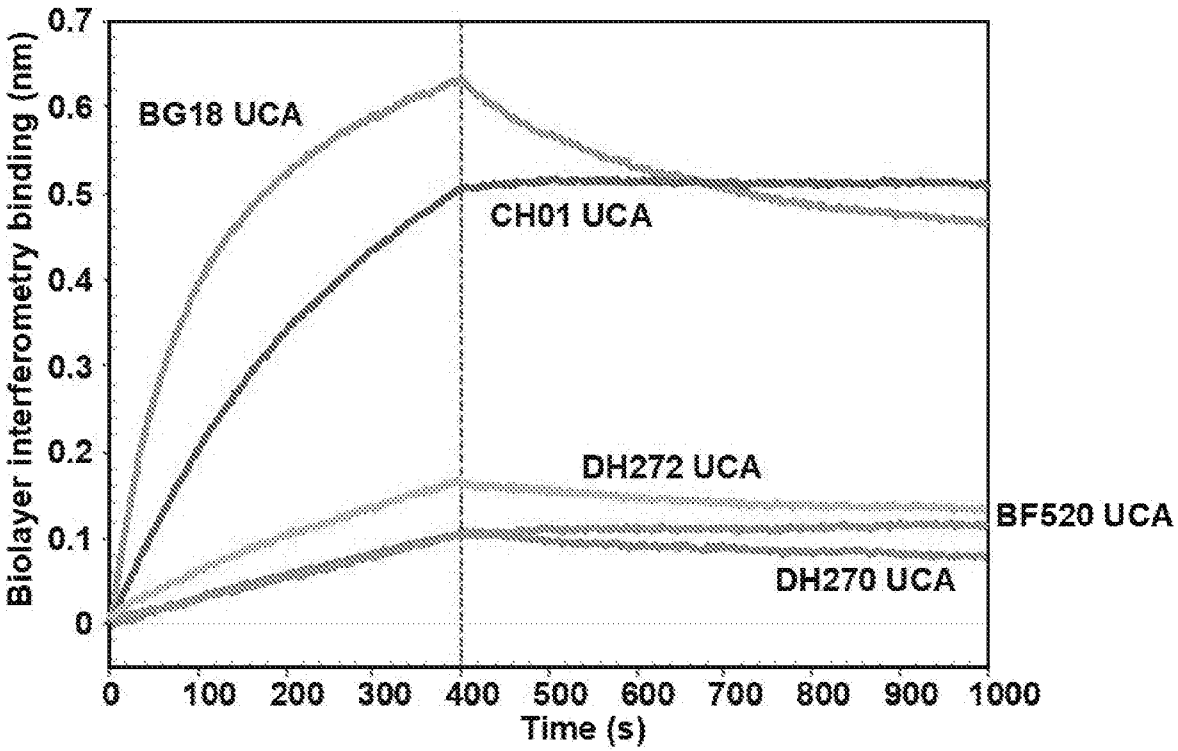

Biolayer interferometry binding (nm)
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
BG18 UCA
CH01 UCA
DH272 UCA
BF520 UCA
DH270 UCA
0
100
200
300
400
500
600
700
800
900
1000
Time (s)

Figure 22

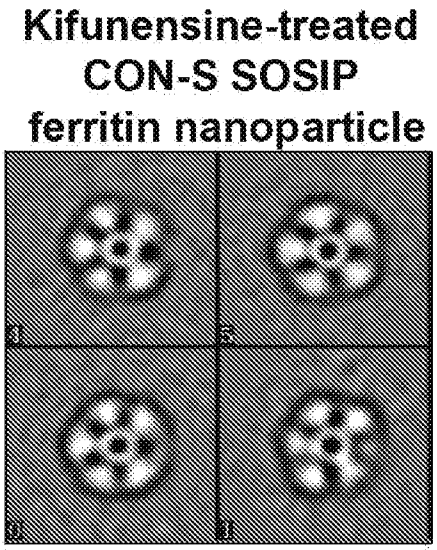


Figure 23A

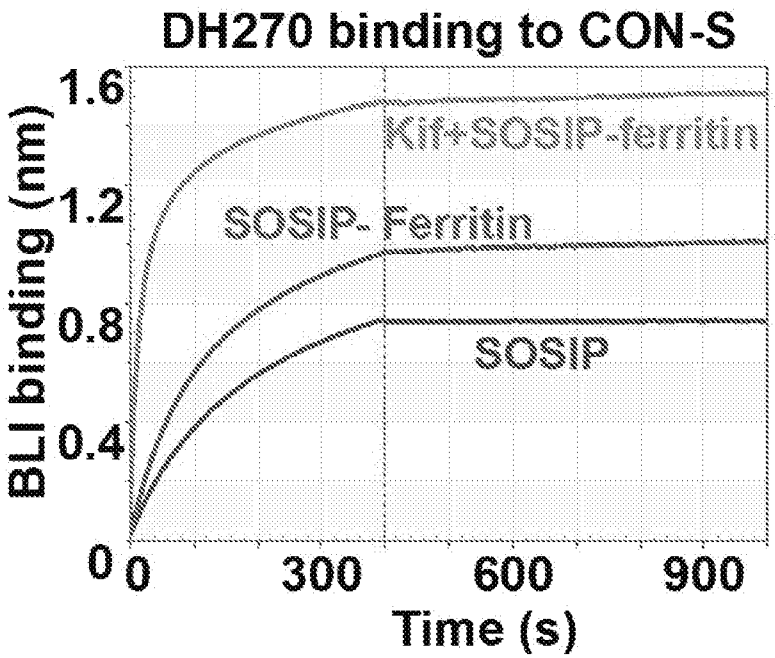


Figure 23B

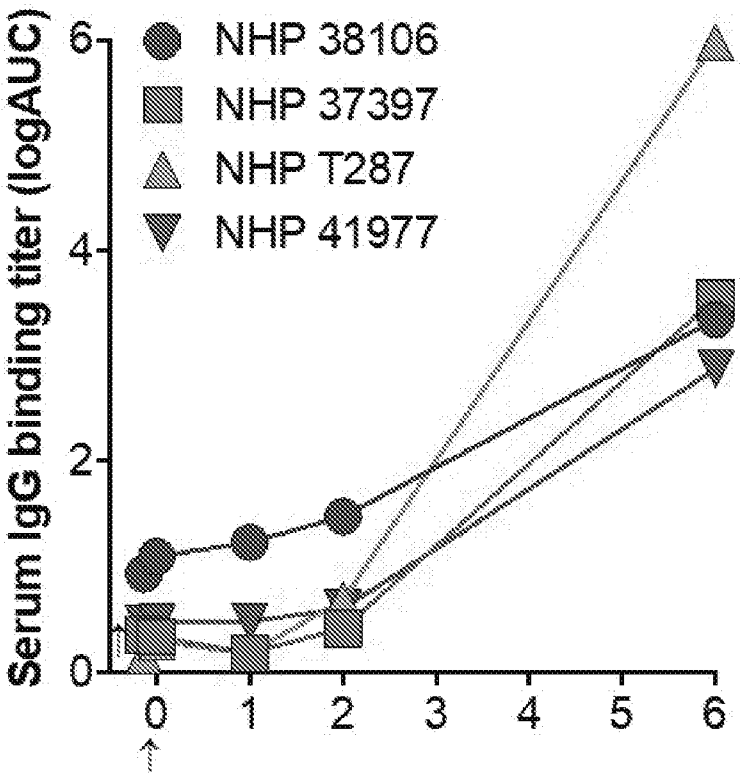


Figure 24A

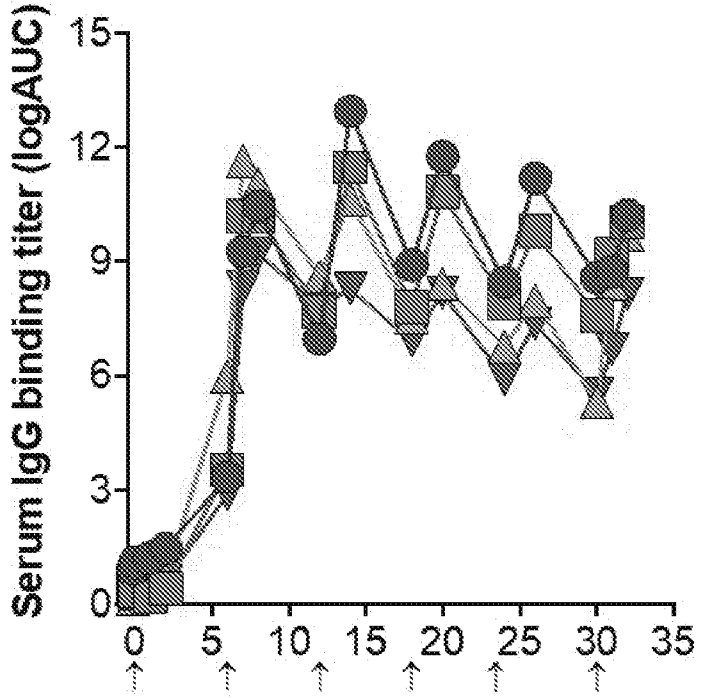


Figure 24B

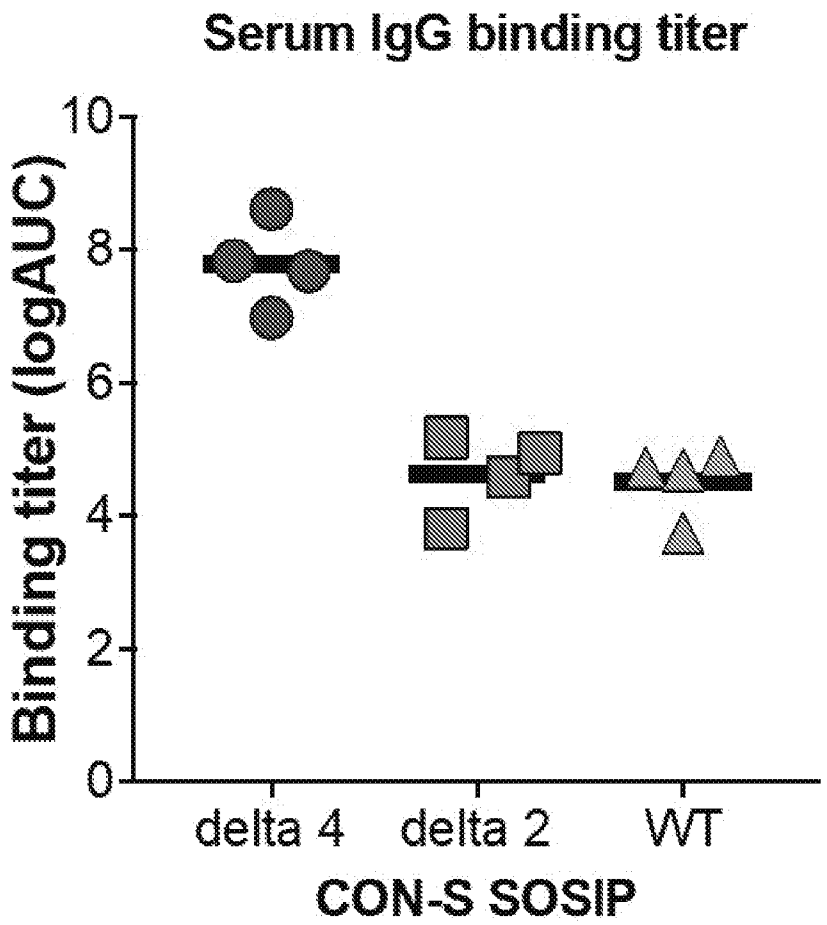

Serum IgG binding titer
Binding titer (logAUC)
10
8
6
4
2
0
delta 4
delta 2
WT
CON-S SOSIP

Figure 25

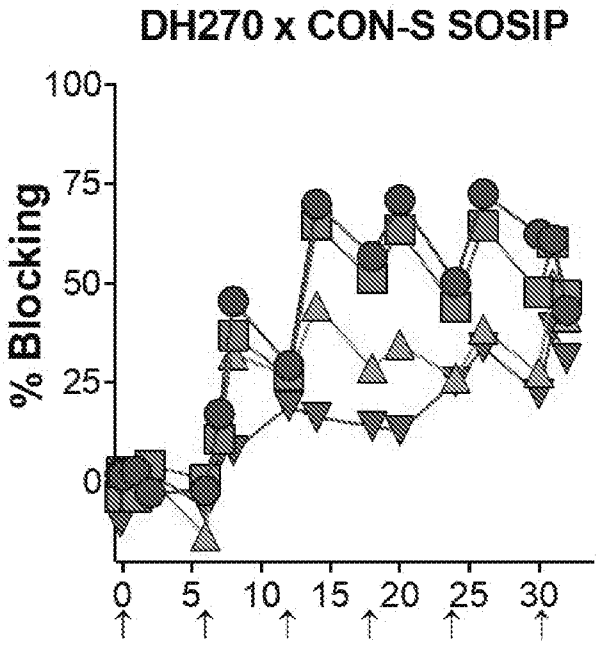


Figure 26A

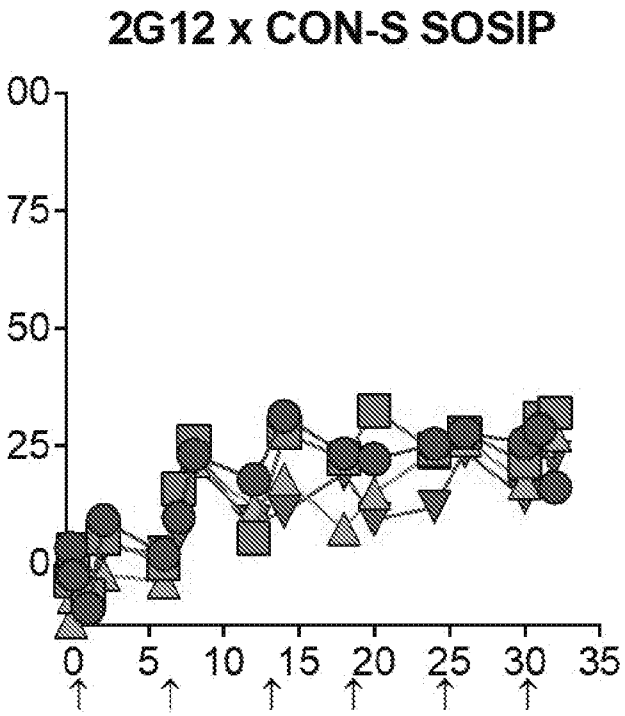


Figure 26B

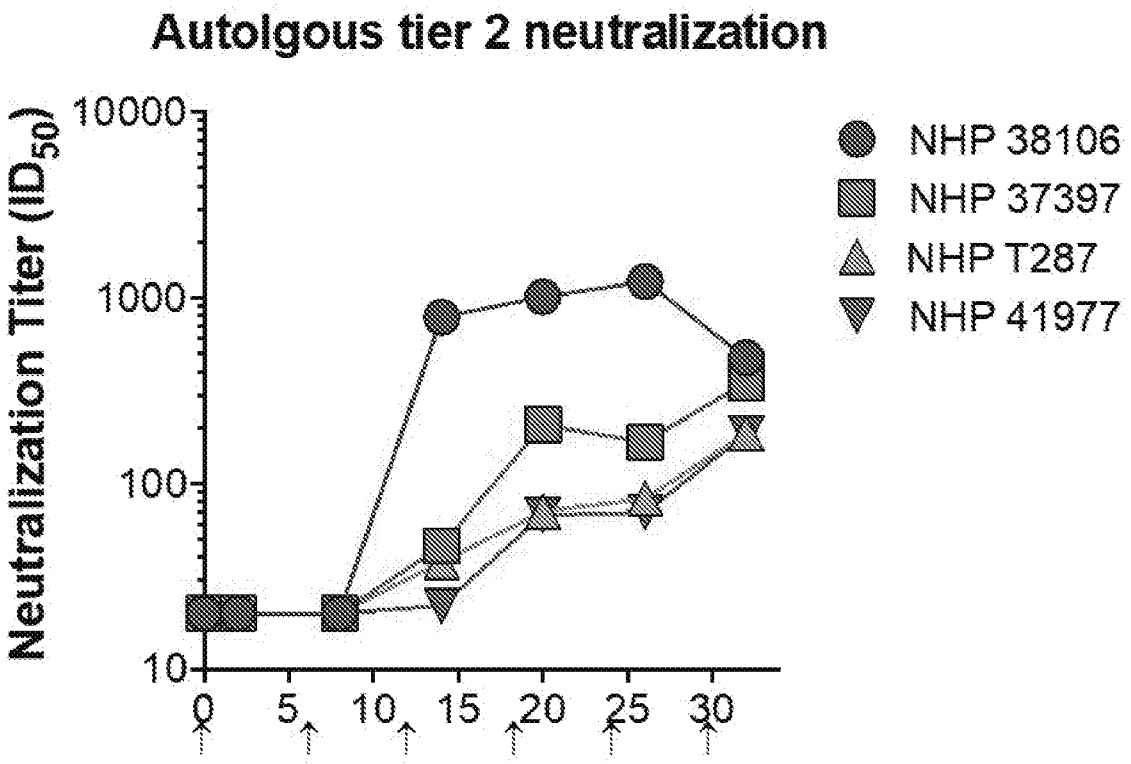

Autolgous tier 2 neutralization
Neutralization Titer (ID_{50})
10000
1000
100
10
0
5
10
15
20
25
30
NHP 38106
NHP 37397
NHP T287
NHP 41977

Figure 27

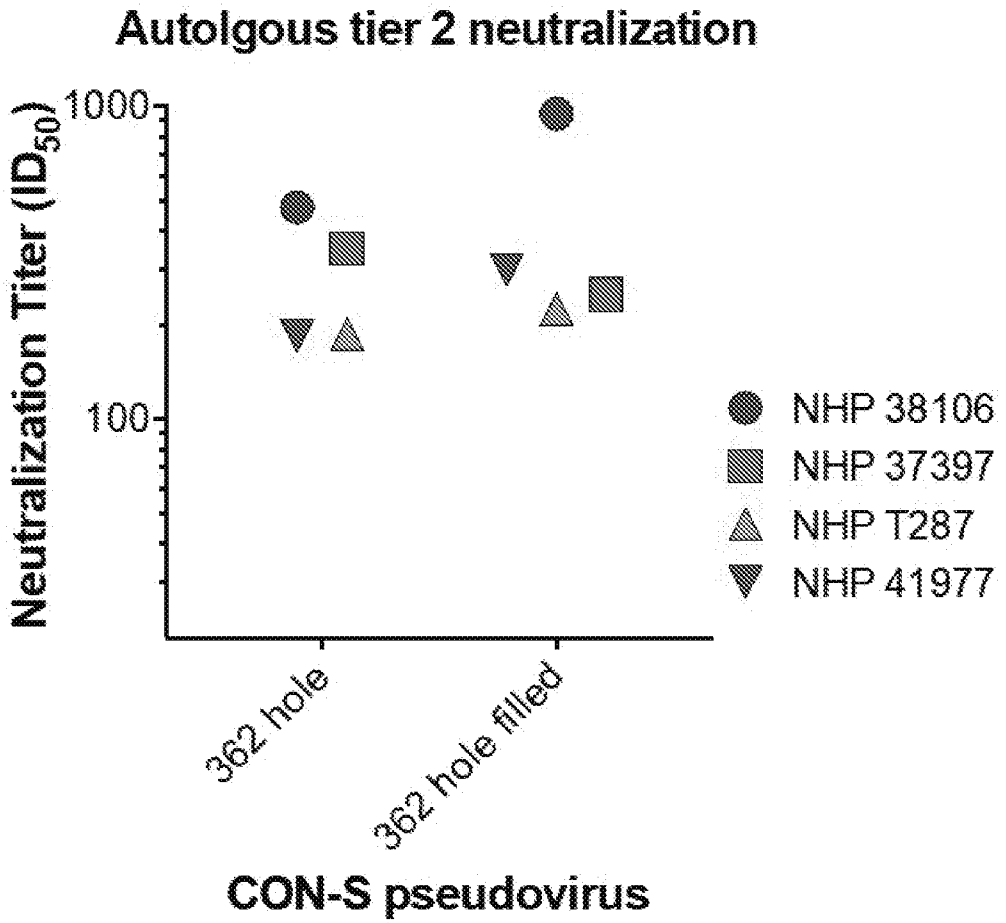

Autolgous tier 2 neutralization
Neutralization Titer (ID_{50})
1000
100
NHP 38106
NHP 37397
NHP T287
NHP 41977
362 hole
362 hole filled
CON-S pseudovirus

Figure 28

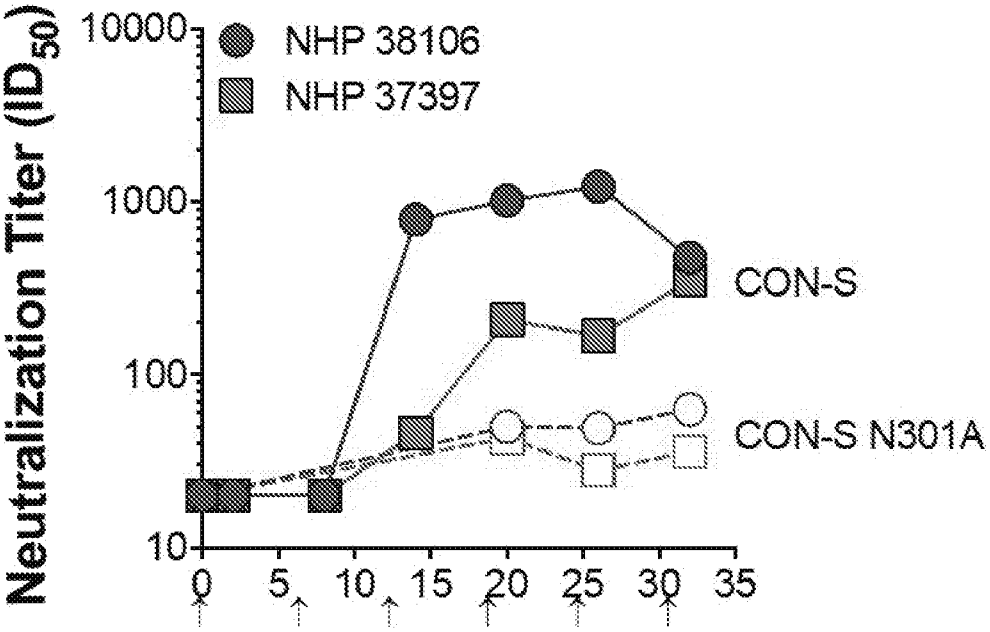


Figure 29A

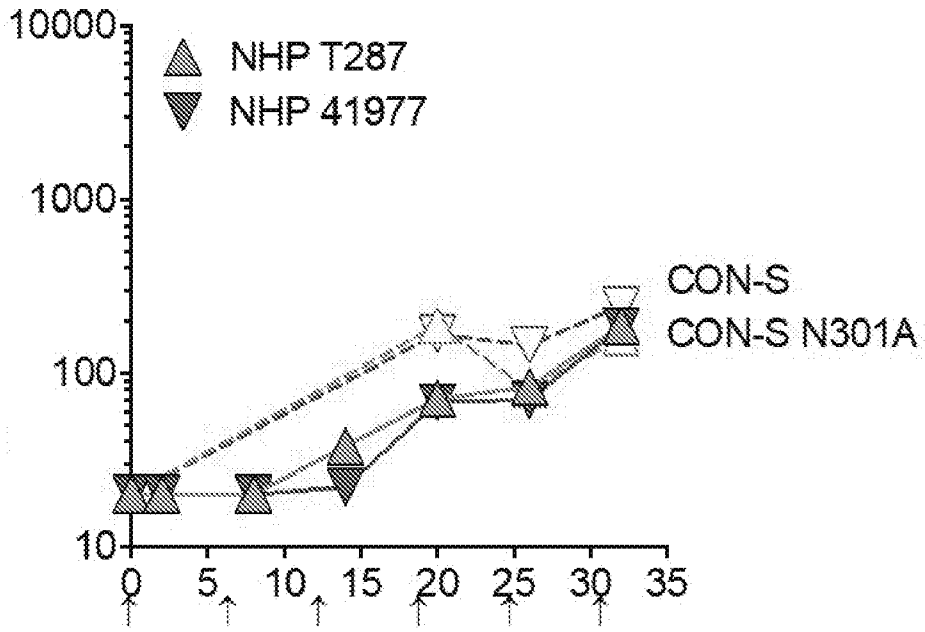


Figure 29B

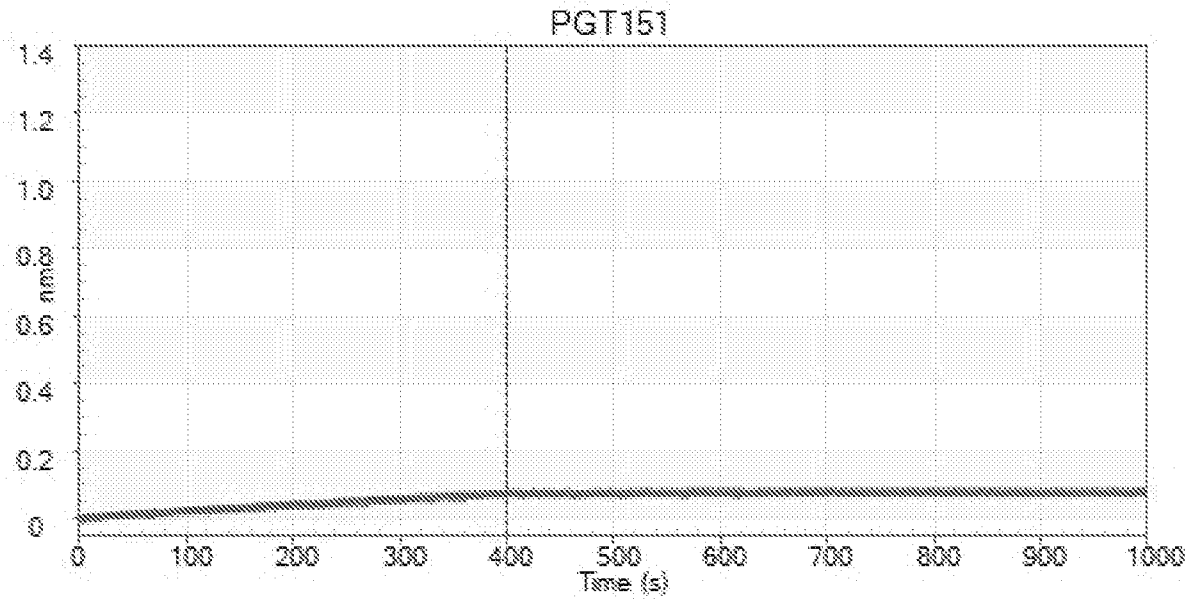


Figure 30A

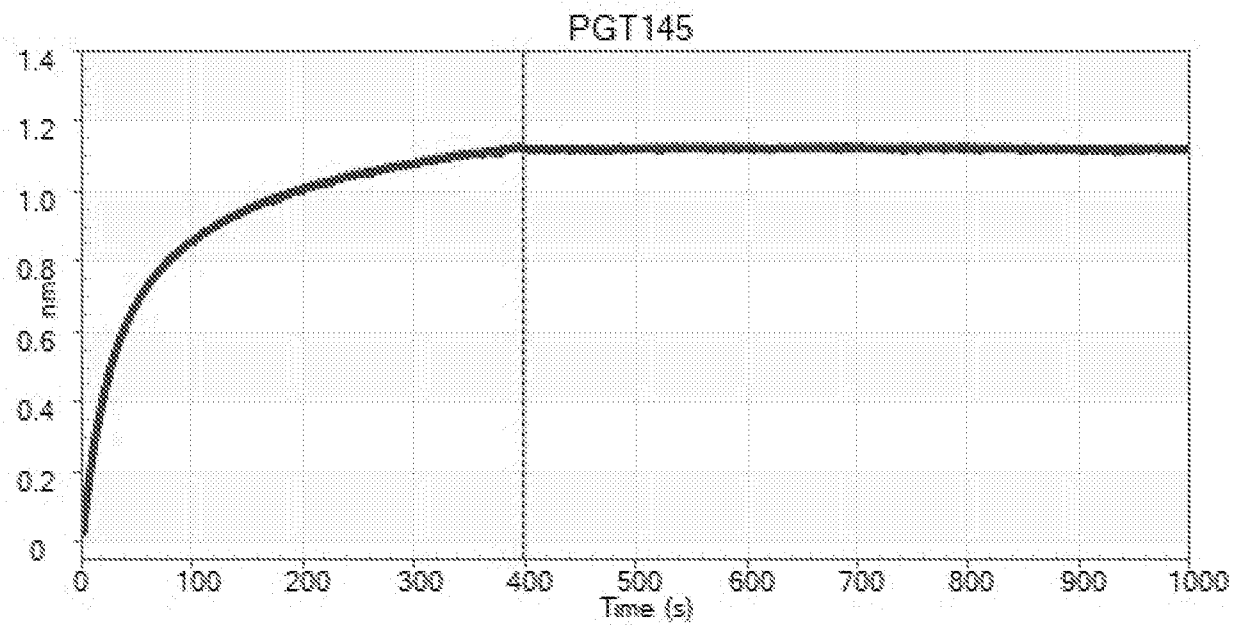


Figure 30B

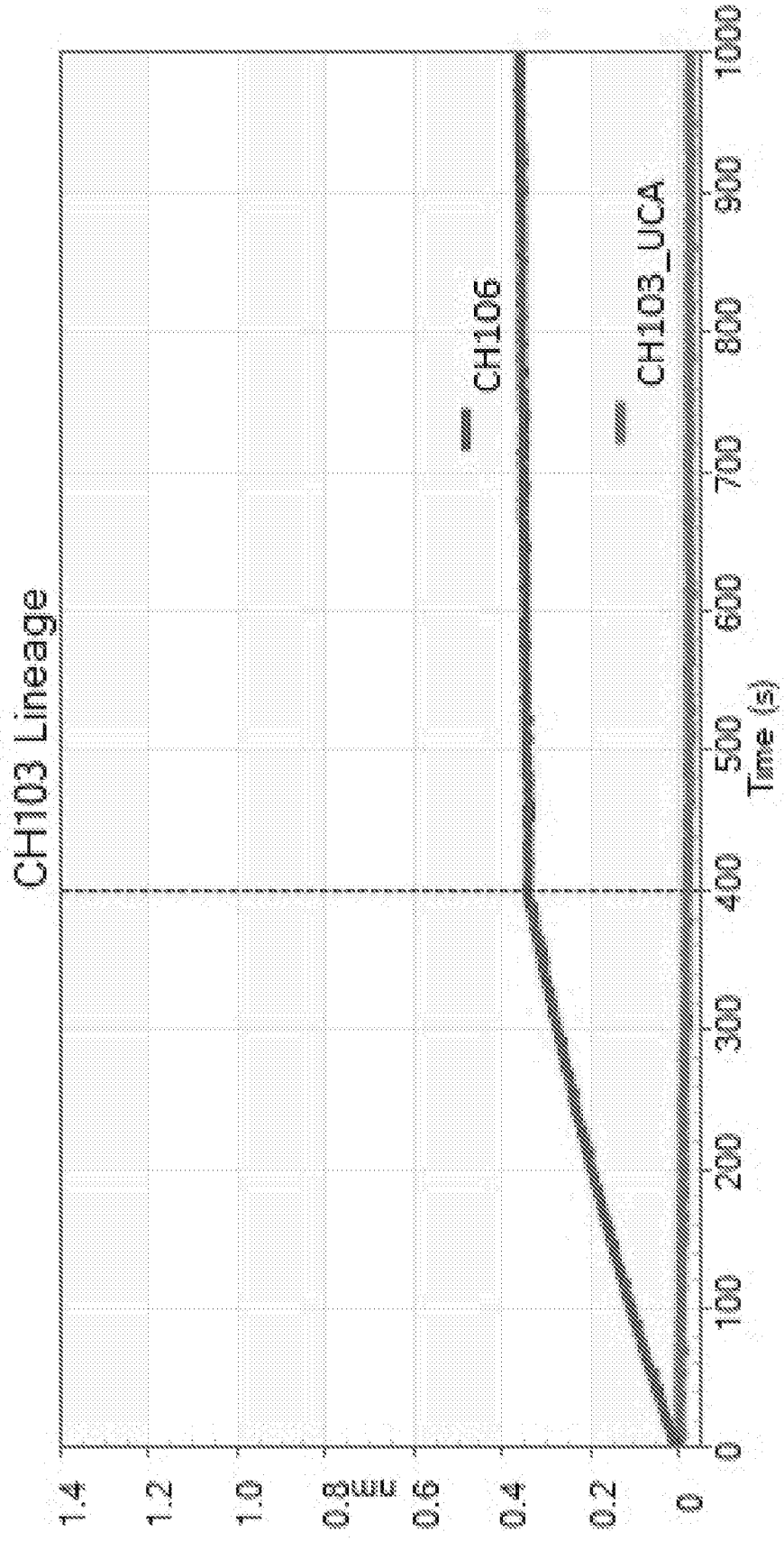

CH103 Lineage
CH106
CH103_UCA
nm
Time (s)
0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
0
100
200
300
400
500
600
700
800
900
1000

Figure 30C

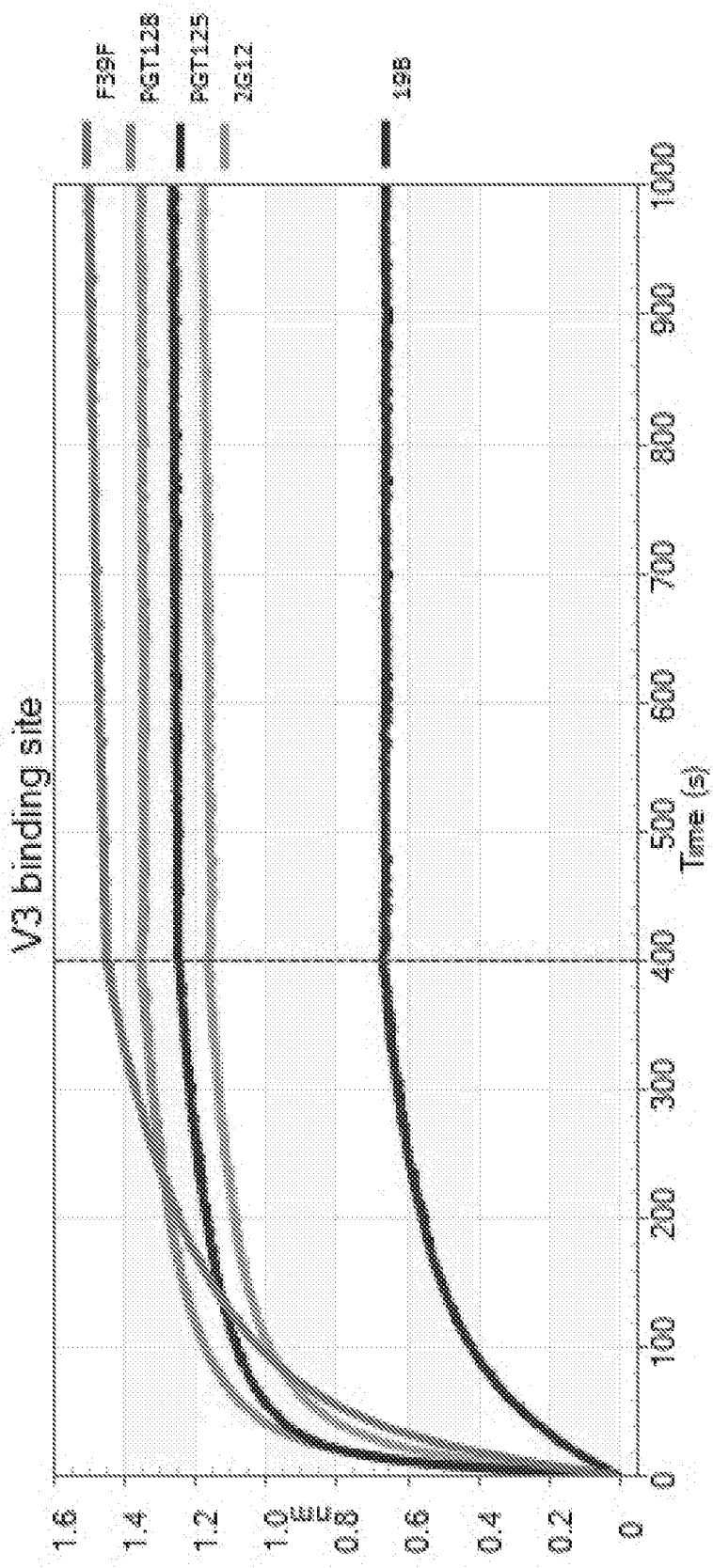

V3 binding site
F39F
PGT128
PGT125
2G12
19b
nm
Time (s)
0
100
200
300
400
500
600
700
800
900
1000
0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6

Figure 30D

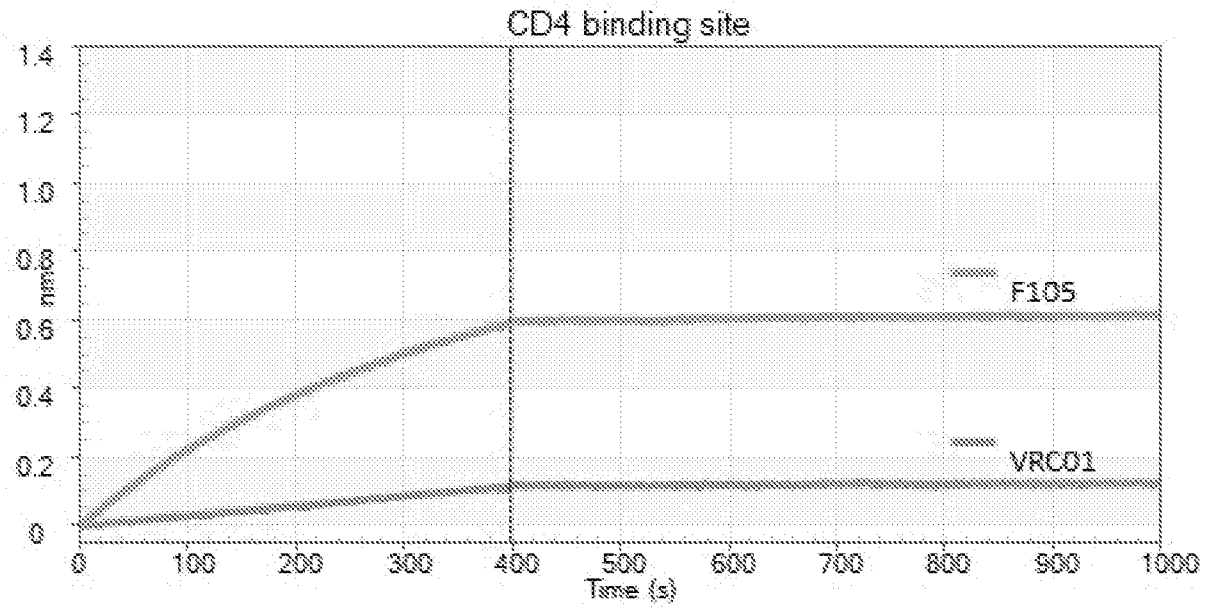


Figure 30E

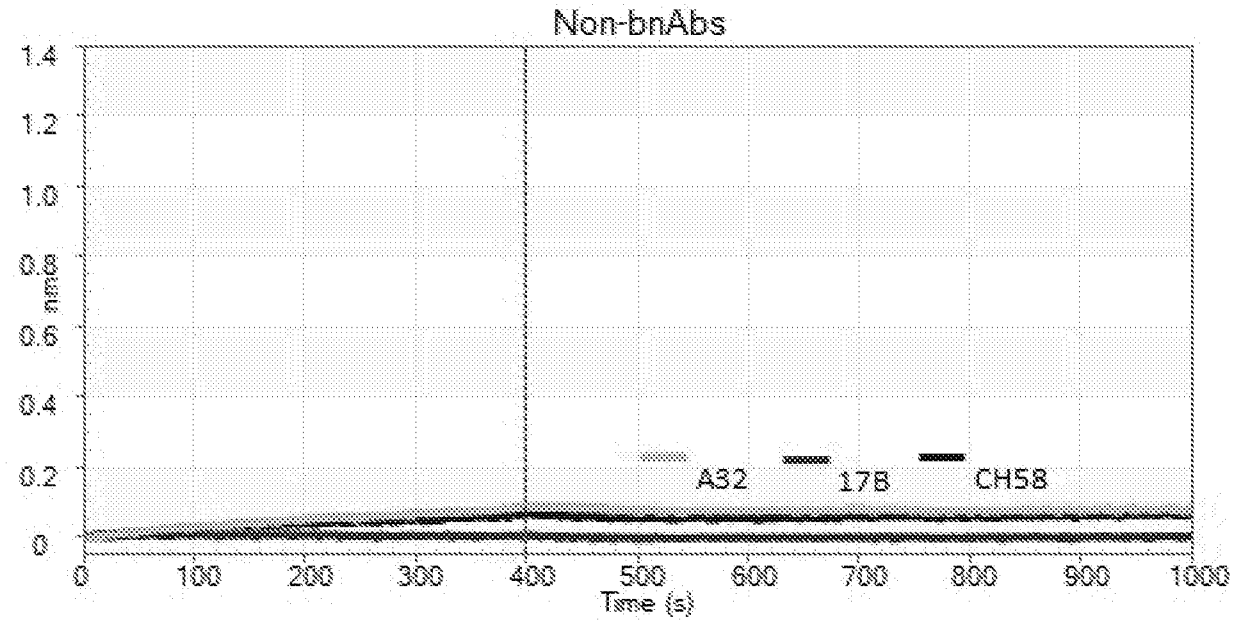


Figure 30F

Loading Sample ID	Response (End of Association)	CH505w100.6R.SOSIP.664v4.1/293F (Control End of Association)
Cat_PGT151_4A B#2 Aliq. 11/20/2017 ESD	0.08	1.08
Cat_PGT145 (3682) Aliq. 11.06.2017 EMC	1.13	0.60
PGT128_LL/293i 92RKK 2.Jul.2015 Aliq. 9.25.2017 EMC	1.36	0.69
PGT125_LL/293i 474HC 8-Apr-2016	1.24	0.52
Cat_2G12_AAA(3362) Aliq. 11/16/2017 ESD	1.17	0.55
VRC01 Purif. IgG Lot: AB001036 Aliq. 5.02.2018 EMC	0.11	0.15
CH103_UCA_4A/293i (Tube A) 388HC (A) 26-Jun-2015	0.00	0.00
Cat CH106 4A (3541) Aliq. 4.21.2015 SJA	0.34	0.34
17B/293i 395AMS 14-Dec-2017	0.00	0.01
19B CL2 Aliq. 8/29/2017 SJA	0.67	0.04
F39F 11/17/06 Aliq. 2/9/12	1.45	0.00
Cat_CH58_4A (3453) Aliq. 4/21/2015 SJA	0.06	0.01
Cat_A32_AAA (3358) Aliq. 7/6/2015 SJA	0.08	0.02
F105/293i 5-Jun-2015 374HC Aliq. 10-Jul-2018	0.59	0.03

Figure 31

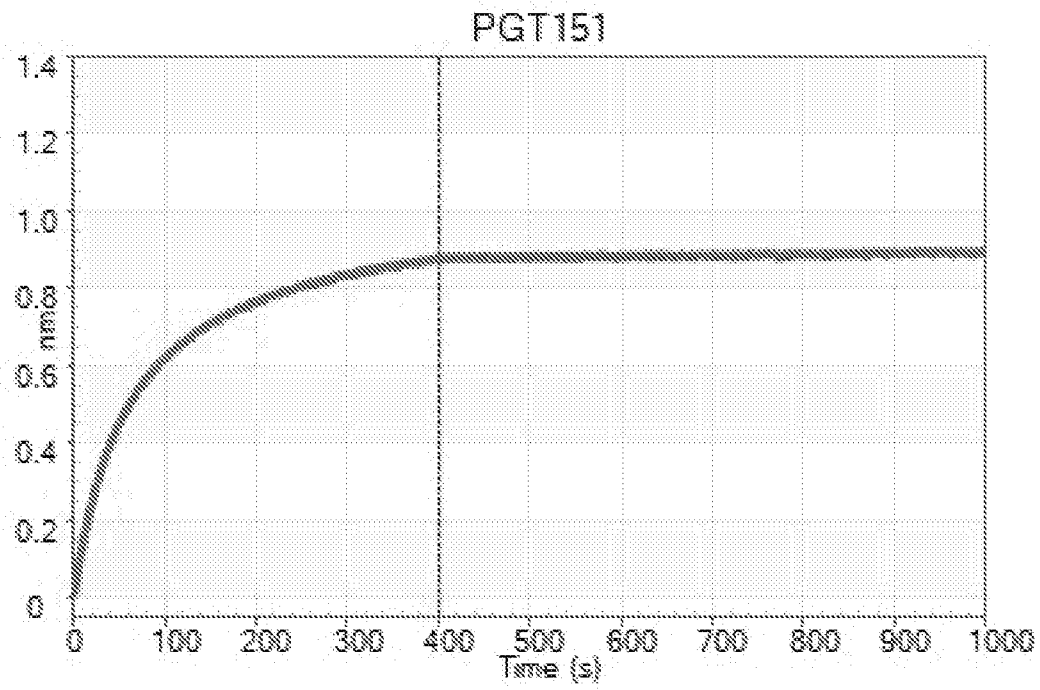


Figure 32A

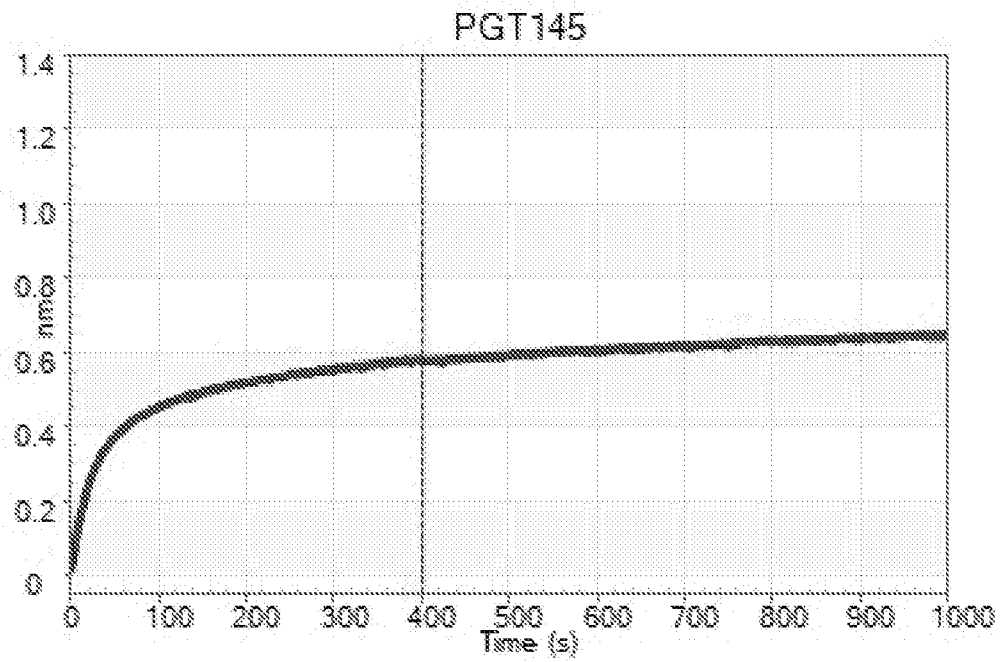


Figure 32B

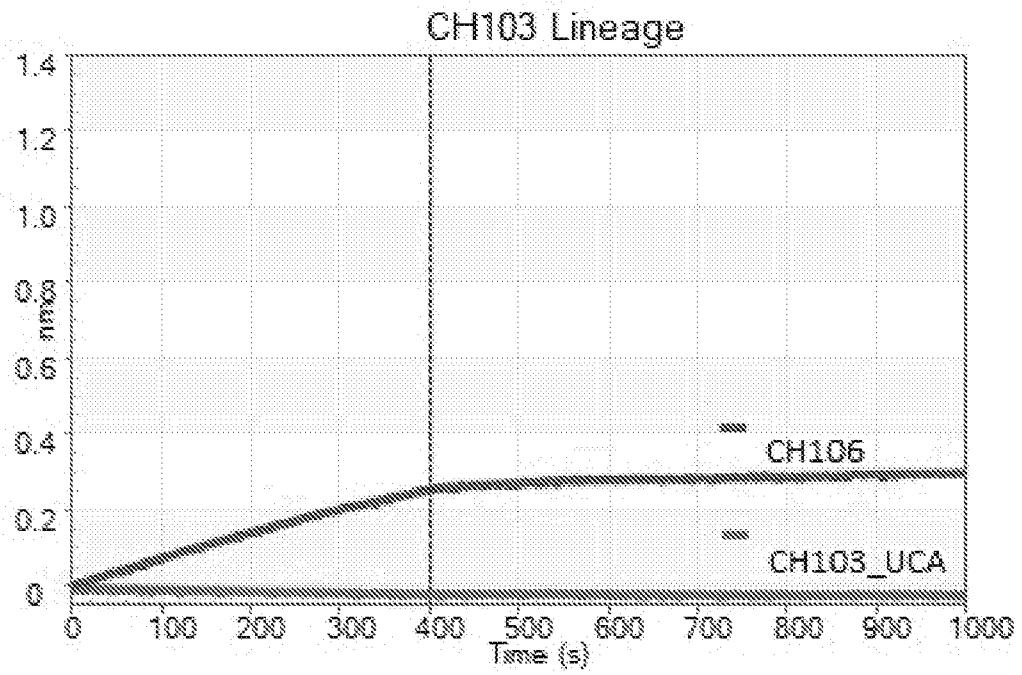


Figure 32C

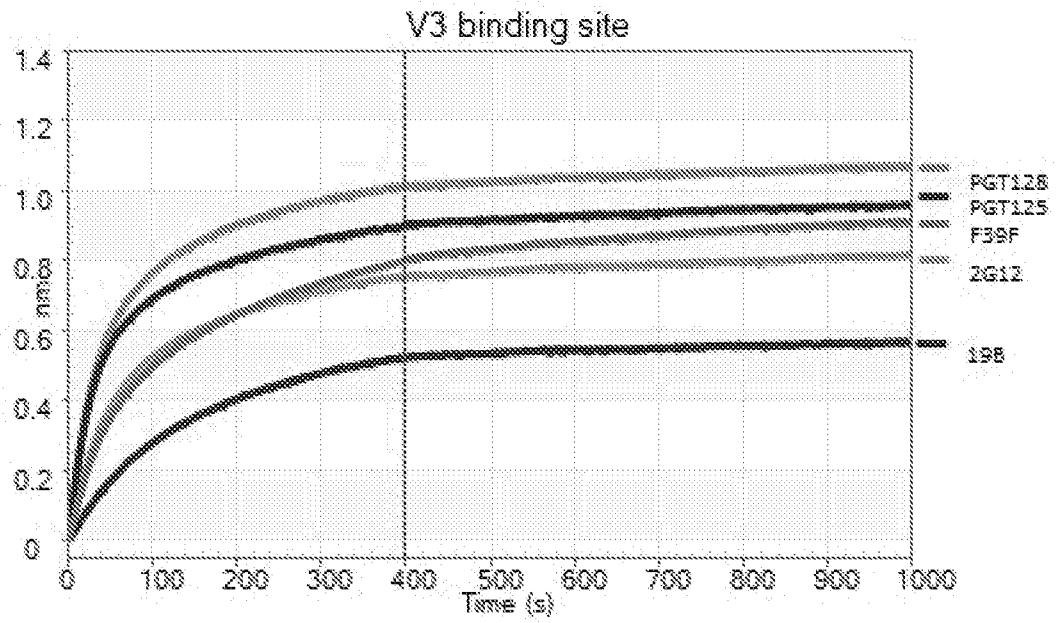


Figure 32D

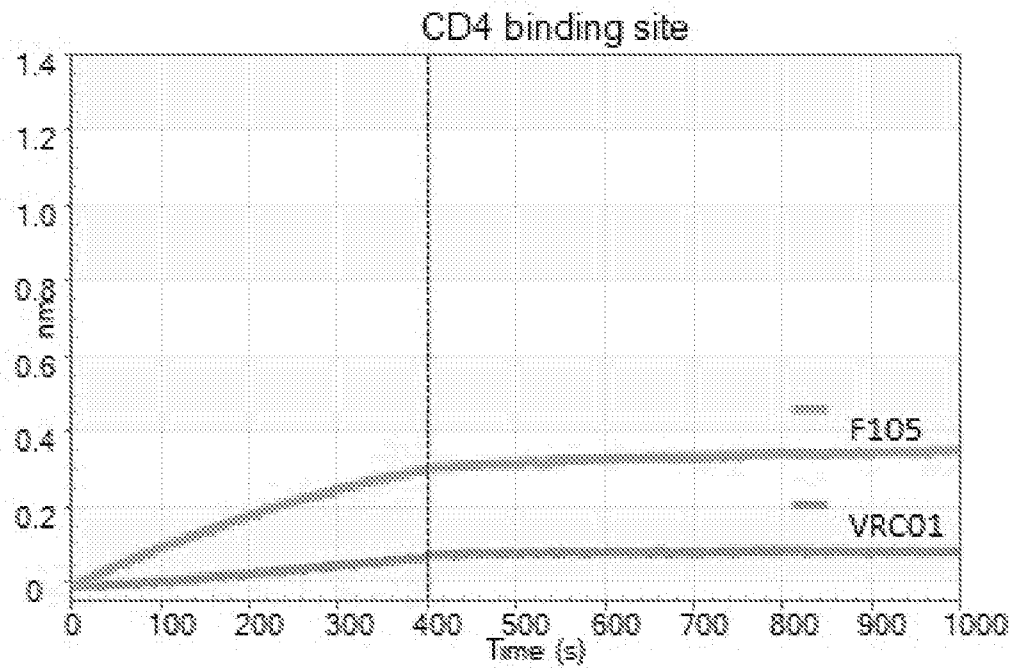


Figure 32E

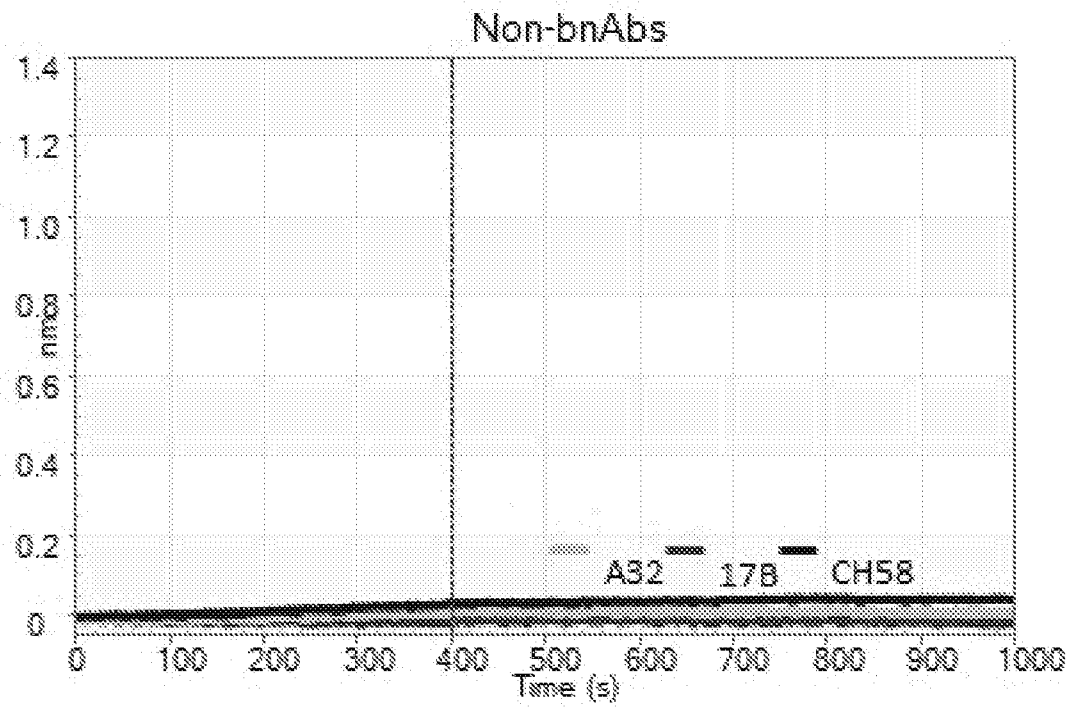


Figure 32F

Loading Sample ID	Response (End of Association)	CH505w100.6R.SOSIP.664v4.1/293F (Control End of Association)
Cat_PGT151_4A B#2 Aliq. 11/20/2017 ESD	0.87	1.06
Cat_PGT145 (3682) Aliq. 11.06.2017 EMC	0.58	0.64
PGT128_LL/293i 92RKK 2.Jul.2015 Aliq. 9.25.2017 EMC	1.01	0.72
PGT125_LL/293i 474HC 8-Apr-2016	0.89	0.57
Cat_2G12_AAA(3362) Aliq. 11/16/2017 ESD	0.75	0.53
VRC01 Purif. IgG Lot: AB001036 Aliq. 5.02.2018 EMC	0.06	0.17
CH103_UCA_4A/293i (Tube A) 388HC (A) 26-Jun-2015	-0.02	0.00
Cat CH106 4A (3541) Aliq. 4.21.2015 SJA	0.25	0.49
17B/293i 395AMS 14-Dec-2017	-0.01	0.00
19B CL2 Aliq. 8/29/2017 SJA	0.52	0.06
F39F-3.5B Aliq. 8/7/14 BWL	0.80	0.00
Cat_CH58_4A (3453) Aliq. 4/21/2015 SJA	0.03	0.00
Cat_A32_AAA (3358) Aliq. 7/6/2015 SJA	0.01	0.01
F105/293i 5-Jun-2015 374HC Aliq. 10-Jul-2018	0.30	0.02

Figure 33

Antigenic profile of CON-S envelopes with various V1 glycosylation sites removed.							
Antibody	BnAb	Epitope	CON-Schim.6R.DS.SOSIP.664_OPT_N130D_N135K_N138S_N141S_ferritin/5uM-Kif	CON-Schim.6R.DS.SOSIP.664_OPT_N130D_N135K_N138S_N141S_ferritin	CON-Schim.6R.DS.SOSIP.664_OPT_N130D_N135K_N138S_N141S	CON-Schim.6R.DS.SOSIP.664_OPT_N138S_N141S	CON-Schim.6R.DS.SOSIP.664_OPT
PGT151	Yes	gp120-gp41 interface	0.08	0.87	1.37	1.15	1.41
PGT145	Yes	V1V2 glycan	1.13	0.58	1.26	0.74	0.91
PGT128	Yes	V3-glycan	1.36	1.01	1.09	0.81	1.04
PGT125	Yes	V3-glycan	1.24	0.89	1.04	0.85	0.99
2G12	Yes	OD glycan	1.17	0.75	0.80	0.71	0.79
VRC01	Yes	CD4 binding site	0.11	0.06	0.09	0.07	0.10
CH106	Yes	CD4 binding site	0.34	0.25	0.06	0.02	0.07
CH103 UCA	No	CD4 binding site	0.00	-0.02	0.00	0.00	-0.01
17B	No	coreceptor binding site	0.00	-0.01	0.00	0.01	0.00
19B	No	V3 peptide	0.67	0.52	0.11	0.05	0.14
F39F	No	V3 peptide	1.45	0.80	0.16	0.09	0.21
CH58	No	V2 peptide	0.06	0.03	0.01	0.01	0.01
A32	No	C1	0.08	0.01	0.02	0.01	0.03
F105	No	CD4 binding site	0.59	0.30	0.01	0.02	0.02
BLI binding of HIV-1 antibodies to CON-S envelopes shows they are antigenic for bnAbs and only weakly antigenic for non-neutralizing antibodies. The lack of 17B binding shows the Envs are not triggered into the CD4-bound conformation.							

Figure 34

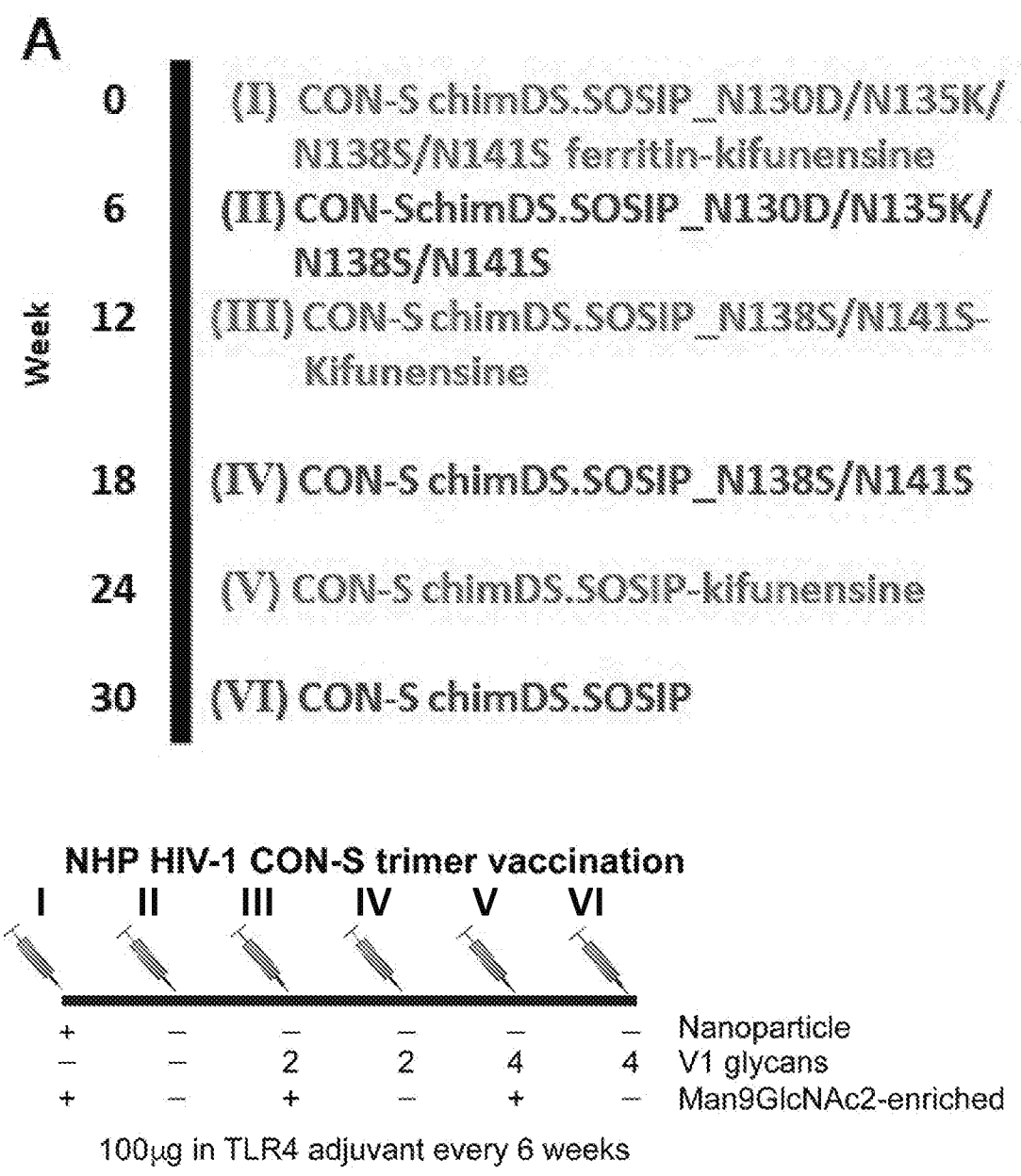

A
Week
0
6
12
18
24
30
(I) CON-S chimDS.SOSIP_N130D/N135K/ N138S/N141S ferritin-kifunensine
(II) CON-SchimDS.SOSIP_N130D/N135K/ N138S/N141S
(III) CON-S chimDS.SOSIP_N138S/N141S-Kifunensine
(IV) CON-S chimDS.SOSIP_N138S/N141S
(V) CON-S chimDS.SOSIP-kifunensine
(VI) CON-S chimDS.SOSIP
NHP HIV-1 CON-S trimer vaccination
I II III IV V VI
+ – – – – – Nanoparticle
– – 2 2 4 4 V1 glycans
+ – + – + – Man9GlcNAc2-enriched
100μg in TLR4 adjuvant every 6 weeks

Figure 3A