



(51) International Patent Classification:

A61K 39/21 (2006.01) *A61K 45/06* (2006.01)
A61K 39/40 (2006.01) *A61P 31/18* (2006.01)
A61K 39/42 (2006.01) *C07K 16/10* (2006.01)

(21) International Application Number:

PCT/US2017/023854

(22) International Filing Date:

23 March 2017 (23.03.2017)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/312,190 23 March 2016 (23.03.2016) US
 62/384,762 8 September 2016 (08.09.2016) US

(71) Applicants: **INTERNATIONAL AIDS VACCINE INITIATIVE** [US/US]; 125 Broad Street, 9th Floor, New York, NY 10004 (US). **THE SCRIPPS RESEARCH INSTITUTE** [US/US]; 10550 North Torrey Pines Road, La Jolla, CA 92037 (US).

(72) Inventors: **STEICHE, Jon**; c/o The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037 (US). **KULP, Dan**; c/o International Aids Vaccine Initiative, 125 Broad Street, 9th Floor, New York, NY 10004 (US). **HU, Xiaozhen**; c/o The Scripps Research In-

stitute, 10550 North Torrey Pines Road, La Jolla, CA 92037 (US). **MENIS, Sergey**; c/o International Aids Vaccine Initiative, 125 Broad Street, 9th Floor, New York, NY 10004 (US). **SCHIEF, William**; c/o The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037 (US). **RAEMISCH, Sebastian**; c/o The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037 (US).

(74) Agents: **KOWALSKI, Thomas, J.** et al.; Vedder Price P.C., 1633 Broadway, New York, NY 10019 (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, 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,

[Continued on next page]

(54) Title: IMMUNOGENIC TRIMERS

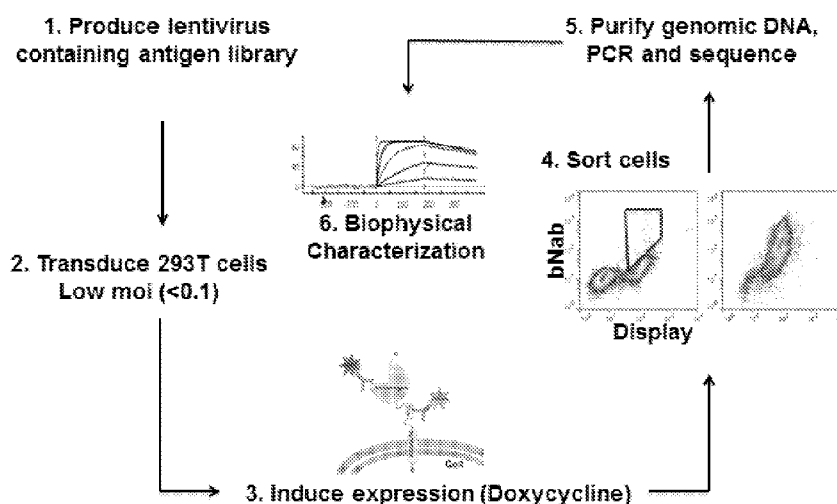


FIG. 1

(57) Abstract: The invention relates to PGT121-germline-targeting designs, trimer stabilization designs, combinations of those two, trimers designed with modified surfaces helpful for immunization regimens, other trimer modifications and on development of trimer nanoparticles and methods of making and using the same.



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).

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))*

IMMUNOGENIC TRIMERS

RELATED APPLICATIONS AND INCORPORATION BY REFERENCE

[0001] This application claims benefit of and priority to US provisional patent application Serial No. 62/312,190 filed March 23, 2016 and Serial No. 62/384,762 filed September 8, 2016.

[0002] All documents cited or referenced herein (“herein cited documents”), and all documents cited or referenced in herein cited documents, together with any manufacturer’s instructions, descriptions, product specifications, and product sheets for any products mentioned herein or in any document incorporated by reference herein, are hereby incorporated herein by reference, and may be employed in the practice of the invention. More specifically, all referenced documents are incorporated by reference to the same extent as if each individual document was specifically and individually indicated to be incorporated by reference.

FIELD OF THE INVENTION

[0003] The invention relates to PGT121-germline-targeting designs, trimer stabilization designs, combinations of those two, trimers designed with modified surfaces helpful for immunization regimens, other trimer modifications and on development of trimer nanoparticles and membrane-bound trimers.

FEDERAL FUNDING LEGEND

[0004] This invention was made with government support under Grant Nos. CHAVI-ID IUM1AI100663 and R01 AI084817 awarded by the National Institute of Allergy and Infectious Disease. This invention was also made with government support under Grant No. P41GM103393 awarded by the National Institutes of Health, National Institute of General Medical Sciences. The government has certain rights in the invention.

BACKGROUND OF THE INVENTION

[0005] AIDS, or Acquired Immunodeficiency Syndrome, is caused by human immunodeficiency virus (HIV) and is characterized by several clinical features including wasting syndromes, central nervous system degeneration and profound immunosuppression that results in opportunistic infections and malignancies. HIV is a member of the lentivirus family of animal retroviruses, which include the visna virus of sheep and the bovine, feline, and simian immunodeficiency viruses (SIV). Two closely related types of HIV, designated

HIV-1 and HIV-2, have been identified thus far, of which HIV-1 is by far the most common cause of AIDS. However, HIV-2, which differs in genomic structure and antigenicity, causes a similar clinical syndrome.

[0006] An infectious HIV particle consists of two identical strands of RNA, each approximately 9.2 kb long, packaged within a core of viral proteins. This core structure is surrounded by a phospholipid bilayer envelope derived from the host cell membrane that also includes virally encoded membrane proteins (Abbas et al., Cellular and Molecular Immunology, 4th edition, W.B. Saunders Company, 2000, p. 454). The HIV genome has the characteristic 5'-LTR-Gag-Pol-Env-LTR-3' organization of the retrovirus family. Long terminal repeats (LTRs) at each end of the viral genome serve as binding sites for transcriptional regulatory proteins from the host and regulate viral integration into the host genome, viral gene expression, and viral replication.

[0007] The HIV genome encodes several structural proteins. The gag gene encodes structural proteins of the nucleocapsid core and matrix. The pol gene encodes reverse transcriptase (RT), integrase (IN), and viral protease (PR) enzymes required for viral replication. The tat gene encodes a protein that is required for elongation of viral transcripts. The rev gene encodes a protein that promotes the nuclear export of incompletely spliced or unspliced viral RNAs. The vif gene product enhances the infectivity of viral particles. The vpr gene product promotes the nuclear import of viral DNA and regulates G2 cell cycle arrest. The vpu and nef genes encode proteins that down regulate host cell CD4 expression and enhance release of virus from infected cells. The env gene encodes the viral envelope glycoprotein that is translated as a 160-kilodalton (kDa) precursor (gp160) and cleaved by a cellular protease to yield the external 120-kDa envelope glycoprotein (gp120) and the transmembrane 41-kDa envelope glycoprotein (gp41), which are required for the infection of cells (Abbas et al., Cellular and Molecular Immunology, 4th edition, W.B. Saunders Company, 2000, pp. 454-456). gp140 is a modified form of the Env glycoprotein, which contains the external 120-kDa envelope glycoprotein portion and the extracellular part of the gp41 portion of Env and has characteristics of both gp120 and gp41. The nef gene is conserved among primate lentiviruses and is one of the first viral genes that is transcribed following infection. In vitro, several functions have been described, including down-regulation of CD4 and MHC class I surface expression, altered T-cell signaling and activation, and enhanced viral infectivity.

[0008] HIV infection initiates with gp120 on the viral particle binding to the CD4 and chemokine receptor molecules (e.g., CXCR4, CCR5) on the cell membrane of target cells such as CD4⁺ T-cells, macrophages and dendritic cells. The bound virus fuses with the target cell and reverse transcribes the RNA genome. The resulting viral DNA integrates into the cellular genome, where it directs the production of new viral RNA, and thereby viral proteins and new virions. These virions bud from the infected cell membrane and establish productive infections in other cells. This process also kills the originally infected cell. HIV can also kill cells indirectly because the CD4 receptor on uninfected T-cells has a strong affinity for gp120 expressed on the surface of infected cells. In this case, the uninfected cells bind, via the CD4 receptor-gp120 interaction, to infected cells and fuse to form a syncytium, which cannot survive. Destruction of CD4⁺ T-lymphocytes, which are critical to immune defense, is a major cause of the progressive immune dysfunction that is the hallmark of AIDS disease progression. The loss of CD4⁺ T cells seriously impairs the body's ability to fight most invaders, but it has a particularly severe impact on the defenses against viruses, fungi, parasites and certain bacteria, including mycobacteria.

[0009] Viruses have evolved a variety of mechanisms to escape antibody recognition, many of which involve features of the viral surface proteins, such as high variability, steric occlusion, and glycan coating. For HIV, the dense shield of glycans that decorate the viral Env protein was once believed to be refractory to antibody recognition, shielding conserved protein epitopes of important functional significance whose greater exposure would result in increased susceptibility to antibody neutralization.

[0010] Citation or identification of any document in this application is not an admission that such document is available as prior art to the present invention.

SUMMARY OF THE INVENTION

[0011] The invention relates to PGT121-germline-targeting designs, trimer stabilization designs, combinations of those two, trimers designed with modified surfaces helpful for immunization regimens, other types of trimer modifications (see, for example, IV: Examples of trimers with combined germline-targeting mutations and stabilization mutations and VI: Additional trimer modifications that add functionality and that can be combined with other types of modifications as described herein) and on development of trimer nanoparticles and membrane-bound trimers. The invention also encompasses combinations of any of the herein described modifications, such as but not limited to,

combinations of stabilization and modified surfaces with nanoparticles or membrane-bound trimers.

[0012] The HIV envelope protein trimer is the target of broadly neutralizing antibodies (bnAbs). The high mannose patch, including the N332-linked glycan at the base of the V3 loop of gp120, is frequently targeted by bnAbs during natural infection and hence is an appealing vaccine epitope. Germline targeting has potential to initiate the elicitation of N332-dependent bnAbs by vaccination, but no immunogen has been reported to bind germline-reverted precursors of N332-dependent bnAbs.

[0013] PGT121 is one of the most broad and potent of the N332-dependent bnAbs. A structural comparison of the mature PGT121 bnAb with its germline-reverted precursors suggested structural regions in BG505 SOSIP that should be modified to generate affinity for the germline. Guided by this structural analysis, Applicants employed mammalian cell surface display to engineer BG505 SOSIP trimer variants with appreciable affinity for germline-reverted PGT121. In the process of developing the germline-targeting immunogen, Applicants produced a series of less mutated trimer immunogens with varying affinities for germline PGT121 and for partially mutated variants of PGT121. Informed by the binding affinities and structural features of the intermediate immunogens, Applicants then postulated sequences of boosting immunogens to guide maturation along different paths following germline activation.

[0014] The invention also encompasses a protein having at least 90% homology or identity with the sequence of the protein of any one of the trimers disclosed herein. The invention also encompasses a protein having at least 95% homology or identity with the sequence of the protein of any one of trimers disclosed herein.

[0015] The invention also encompasses any nucleic acid encoding the protein of any one of the trimers disclosed herein. The invention also encompasses a nucleic acid having at least 90% or 95% homology or identity with the sequence of said nucleic acid.

[0016] The present invention also encompasses methods for eliciting an immune response which may comprise systemically administering to an animal in need thereof an effective amount of the protein of any one of the trimers disclosed herein. The animal may be a mammal, advantageously a human.

[0017] Accordingly, it is an object of the invention not to encompass within the invention any previously known product, process of making the product, or method of using the product such that Applicants reserve the right and hereby disclose a disclaimer of any

previously known product, process, or method. It is further noted that the invention does not intend to encompass within the scope of the invention any product, process, or making of the product or method of using the product, which does not meet the written description and enablement requirements of the USPTO (35 U.S.C. §112, first paragraph) or the EPO (Article 83 of the EPC), such that Applicants reserve the right and hereby disclose a disclaimer of any previously described product, process of making the product, or method of using the product. It may be advantageous in the practice of the invention to be in compliance with Art. 53(c) EPC and Rule 28(b) and (c) EPC. Nothing herein is to be construed as a promise.

[0018] It is noted that in this disclosure and particularly in the claims and/or paragraphs, terms such as "comprises", "comprised", "comprising" and the like can have the meaning attributed to it in U.S. Patent law; e.g., they can mean "includes", "included", "including", and the like; and that terms such as "consisting essentially of" and "consists essentially of" have the meaning ascribed to them in U.S. Patent law, e.g., they allow for elements not explicitly recited, but exclude elements that are found in the prior art or that affect a basic or novel characteristic of the invention.

[0019] These and other embodiments are disclosed or are obvious from and encompassed by, the following Detailed Description.

BRIEF DESCRIPTION OF THE DRAWINGS

[0020] The patent or application file contains at least one drawing executed in color. Copies of this patent or patent application publication with color drawing(s) will be provided by the Office upon request and payment of the necessary fee.

[0021] The following detailed description, given by way of example, but not intended to limit the invention solely to the specific embodiments described, may best be understood in conjunction with the accompanying drawings.

[0022] FIG. 1 depicts a mammalian display/ directed evolution overview.

[0023] FIG. 2 depicts a mammalian display strategy using partially mutated PGT121 intermediate antibodies

[0024] FIG. 3 depicts binding affinities (SPR K_d (nM)) of germline-targeting gp120s (monomeric gp120 used for SPR, but SOSIP used for immunization) NB, no binding; NB**, no binding at 150 uM.

[0025] FIG. 4 depicts biophysical properties of a SOSIP trimer engineered by mammalian display directed evolution and having improved thermostability, yield and antigenic profile. A BG505 SOSIP trimer whole gene saturation mutagenesis library was FACS sorted on the

surface of mammalian cells for reduced binding to B6 and 4025 and enhanced binding to PGT145 and PGT151. Additional combinatorial libraries based on the initial screen as well as an Env sequence alignment were sorted by the same method. This resulted in BG505 SOSIP-MD39 which contained 11 mutations and had improved expression and thermostability while maintaining a native-like antigenic profile. V3 mAb SPR was measured with IgG as the ligand and trimer as the analyte to allow avidity for maximum sensitivity, whereas bnAb Kds were measured as monovalent interactions. These stabilizing mutations were combined with PGT121 germline-targeting mutations to make native-like trimers with enhanced stability that engage PGT121 germline precursors.

[0026] FIG. 5 depicts postulated sequences of boosting immunogens to guide maturation along different paths following germline activation.

[0027] FIG. 6A-B depicts a mammalian display-directed evolution design pathway for PGT121 germline-targeting Env-based immunogens. (FIG. 6A) Models of the PGT121 epitope are shown for each immunogen, with positions of germline-targeting mutations colored red and glycans depicted with cyan spheres. The epitope of BG505 is colored yellow (variable loop 1) and pink (variable loop 3). The paratope of PGT122 is mapped onto the epitope of BG505 and shown in tube representation (heavy chain in blue, light chain purple) (FIG. 6B) Binding K_Ds of mature, intermediately mutated and germline-reverted variants of PGT121 for BG505-gp120 and germline-targeting gp120s, determined by SPR. SPR K_Ds are the average of two or three experiments. *Indicates complex binding kinetics. WB, weak binding; (-), not done.

[0028] FIG. 7A-C depicts a design of mutations to stabilize BG505-SOSIP and germline-targeting native-like trimers. (FIG. 7A) Biophysical properties of stabilized BG505-SOSIP and germline-targeting trimers. Antigenic profile was assessed by SPR, thermostability measurements were made by DSC, and expression was determined as yield of purified protein relative to BG505-SOSIP made with PEI or 293Fectin transfection reagents in 293F cells. Monovalent K_Ds were measured by SPR with trimer ligand and Fab analyte except for PGT145 and PGDM1400, for which monovalent K_Ds were determined with IgG ligand and trimer analyte. For PGT151, a 1-to-1 binding model gave a relatively poor kinetic fit. (FIG. 7B) Antigenic profile of stabilized BG505-SOSIP and germline-targeting trimers by ELISA. Data are representative of two independent experiments, each done in duplicate. (FIG. 7C) Mammalian display-directed evolution design pathways for engineering stabilized native-like trimers.

[0029] FIG. 8A-F depicts sequential boosting schemes employing a native-like trimer cocktail and germline-targeting design intermediates. (FIG. 8A) Side view of a single PGT122 Fab (light blue cartoon and semi-transparent surface) bound to the BG505 SOSIP native-like gp140 trimer, based on PDB ID: 4NCO. The PGT122-bound gp140 subunit is shown in wheat-colored cartoon; the V1, V2, V4, V5 variable loops on that subunit, modeled wherever missing in the crystal structure, are shown in yellow (V1), olive (V2), teal (V4), magenta (V5); the N332 glycan is shown as red spheres; and the two other gp140 subunits are shown as gray surfaces. (FIG. 8B) Same model as in (FIG. 8A), except that glycosylation sites on the trimer have been decorated with Man8GlcNAc2 glycans shown as spheres (V1 glycans, red; V2 glycans, olive; V4 glycans, teal; N156 glycan, magenta; N332 glycan, red; all other glycans, gray), and all trimer subunits are gray surface. (FIG. 8C) Two-dimensional histogram of variable loop (V1, V2, V4, V5) length and number of glycosylation sites among 3,897 unique HIV Env sequences isolated from infected individuals obtained from www.hiv.lanl.gov. Frequency is indicated by the color scale shown for each loop. The length and number of glycosylation sites for each loop of the native-like VLC trimers are indicated. Further elaboration of this cocktail could include accounting for sequence variation at non-variable-loop positions within the N332-epitope region (Figure 14C). (FIG. 8D) Basic scheme in which a germline-targeting prime (10MUT or 11MUT_B) is boosted by a native-like trimer (BG505) and then by a cocktail of native-like trimers. (FIG. 8E) Diagram illustrating seven boosting schemes employing germline-targeting design intermediates (7MUT, 6MUT, 5MUT, 3MUT) as boosts after a germline-targeting prime and before a native-like trimer; the scheme in A is included for reference. Relative affinity drops (in fold affinity decrease) for each boost, computed from Figure 6B as described in the text, are indicated as red numbers. (FIG. 8F) Linear diagrams of three of the best boosting schemes as ranked by favoring those with the smallest maximum affinity drop.

[0030] FIG. 9A-C depicts mammalian display overview related to Figure 6. (FIG. 9A) Schematic of mammalian display procedure. (FIG. 9B) Example FACS plots for unsorted mammalian display library (top) and the same library sorted 3 times (bottom), see Extended Experimental Procedures. (FIG. 9C) Sequences of BG505-SOSIP and BG505-gp120 used for mammalian display. Leader peptides are shown in red, the cMyc epitope is shown in green and the PDGFR TM is shown in blue.

[0031] FIG. 10 depicts germline-reverted PGT121 Abs, related to Figure 6. Sequences of PGT121 and germline reverted variants with mutations highlighted relative to germline V4-59/D3-3/J6 (heavy chain) and V3-21/J3 (light chain) genes.

[0032] FIG. 11 depicts PGT121 germline targeting gp120s and gp140s, related to Figure 6. Designed germline targeting trimer and gp120 sequences are shown with mutations from BG505-SOSIP highlighted. BG505-SOSIP.D664 contains the mutation T332N, not highlighted.

[0033] FIG. 12A-B depicts SPR binding data for V3 Abs binding to gp140 SOSIPs and their matching gp120s, related to Figure 7. (FIG. 12A) Comparison of V3 binding between WT BG505-SOSIP and BG505-SOSIP variants containing R304V and A319Y V3 mutations. For the 4025 SPR plots, the gp140 concentrations tested are shown in μM next to each relevant sensogram and are equivalent for all V3 Abs tested. (FIG. 12B) Comparison of V3 binding to WT gp120 and gp120-MD16 containing the R304V and A319Y mutations. For the 4025 SPR plots, the gp120 concentrations tested are shown in μM next to each relevant sensogram and are equivalent for all V3 Abs tested.

[0034] FIG. 13A-C depicts trimers with improved thermostability, expression, or antigenic profile, related to Figure 7. (FIG. 13A) Sequences of designed stabilized native-like trimers with mutations from BG505 SOSIP.D664 highlighted. (FIG. 13B) Melting temperature of stabilized trimers as assessed by DSC. (FIG. 13C) Native-like trimers with and without stabilizing mutations were transiently transfected in 293F cells and expression levels were determined by capture ELISA using PGT145 Fab for immobilization and PGT151 IgG for detection. Values are the mean $\pm$ SD of 3 replicate transfections.

[0035] FIG. 14A-C depicts sequences and biophysical properties of the MD39-based VLC native-like trimer cocktail, and N332-epitope sequence diversity. (FIG. 14A) Biophysical characterization of the MD39-based VLC cocktail. (FIG. 14B) Sequences of MD39-based VLC cocktail members, with changes relative to MD39 highlighted in green. (FIG. 14C) List of interface positions on the BG505 SOSIP trimer near N332-supersite bnAb epitopes (PGT122, PGT128, PGT135), showing the frequencies of the amino acids found at those positions in 10% or more of 3,897 unique HIV Env sequences isolated from infected individuals obtained from www.hiv.lanl.gov.

[0036] FIG. 15A-G depicts properties of BG505 Olio6, Olio6 CD4KO, MD39 and MD37. FIG. 15A. Binding dissociation constants (KDs) measured by SPR for BG505 SOSIP, BG505 MD39 and BG505 Olio6 binding to various bnAbs. KDs were measured using trimer

analyte and IgG captured on the sensor as ligand for V1V2 apex antibodies that bind as one Fab per trimer, but for other antibodies the KDs were measured with trimer captured as ligand and Fabs as analytes. FIG. 15B. Binding of the native-like trimers BG505 SOSIP and BG505 Olio6 and the non-native trimer BG505 gp120 foldon to multiple non-nAbs, assessed by SPR. Trimers were analytes and IgG were captured on the sensor chip as ligand. Binding was assessed at 1 μ M trimer analyte concentration and is displayed as the ratio of the observed response units to the response units expected for a 1 to 1 binding interaction between each gp120 on the trimer and each Fab on the IgG. Lower values correspond to less binding. FIG. 15C. Binding of the native-like trimers BG505 SOSIP, BG505 MD39, BG505 Olio6 and BG505 Olio6 CD4KO, and the non-native trimer BG505 gp120 foldon, to various bnAbs assessed by ELISA and displayed as the antibody concentration that gives a half-maximal response (EC₅₀). FIG. 15D. Binding of the native-like trimers BG505 SOSIP, BG505 MD39, BG505 Olio6 and BG505 MD37 to various bnAbs by ELISA and displayed as the area under the curve (AUC). FIG. 15E. Binding of BG505 Olio6 and BG505 Olio6 CD4KO to CD4-IgG assessed by SPR. CD4-IgG was captured on the sensor as ligand and trimers were analytes. No binding was detected for BG505 Olio6 CD4KO at a maximum trimer concentration of 1 μ M. FIG. 15F. Binding of BG505 SOSIP and BG505 Olio6 CD4KO to human CD4⁺ T cells by flow cytometry. Trimers were labelled with two different fluorescent probes, cells were stained with trimers for 30 mins, and trimer-specific binding was detected as the percentage of cells in the double-positive population. 99.1% of cells bound to BG505 SOSIP but only 0.1% of cells bound to BG505 Olio6 CD4KO. FIG. 15G. Size exclusion chromatography coupled inline with multi-angle light scattering (SEC-MALS) analysis showing as examples that, in solution, BG505 MD39, BG505 Olio6, BG505 MD39c, and BG505 MD39 CP1.2 GRSF4 have the molecular weight expected of a trimer.

[0037] FIG. 16A-C depicts biophysical properties of stabilized trimers for BG505 and other HIV strains. FIG. 16A. Binding dissociation constants (KDs) measured by SPR for multiple stabilized trimers binding to different bnAbs (PGDM1400, PGT145, PGT121, PGT128, VRC01, 12A12, 3BNC60 and PGT151). Also, the last column shows the melting temperature (T_m) measured by differential scanning calorimetry (DSC). FIG. 16B. DSC melting data for BG505 MD37, BG505 MD39, BG505 Olio6 and BG505 MD39C. FIG. 16C. DSC melting data for BG505 MD53, BG505 MD39 CP1.1, BG505 MD39 CP1.2 and BG505 MD39 link14.

[0038] FIG. 17A-C depicts properties of BG505 MD39 GRSF4. FIG. 17A. Side and Bottom views of a model of BG505 MD39 GRSF4, illustrating the locations of the five newly introduced glycosylation sites at positions 80, 241, 289, 657 and 665. Addition of these glycosylation sites is intended to mask an otherwise exposed surface on the soluble BG505 trimer, a surface that is not targeted by the classes of bnAbs Applicants aim to elicit against the CD4bs, V2 apex and N332 supersite. Positions 241 and 289 are glycosylation sites that are relatively conserved across HIV strains but are absent in BG505. Positions 80 and 657 are not glycosylation sites in 2869 HIV Env sequences Applicants examined, and 665 appears as a glycosylation site in 2 of 2869 sequences. FIG. 17B. Binding of the native-like trimer BG505 MD39 GRSF4 to various bnAbs and the non-nAbs 17b, 3074, 4025 and B6, assessed by SPR. Trimers were analytes and IgG were captured on the sensor chip as ligand. Binding was assessed at 1 μ M trimer analyte concentration and is displayed as the ratio of the observed response units to the response units expected for a 1 to 1 binding interaction between each gp120 on the trimer and each Fab on the IgG. Lower values correspond to less binding. FIG. 17C. Data illustrating that the additional glycans present in MD39 GRSF4 serve to block undesired responses to soluble trimers. ELISA area under the curve (AUC) values are displayed for six VRC01 gH mice from Applicants' study in Briney et al. Cell 2016. These six mice received four immunizations : eOD-GT8 60mer, core N276D 60mer, and two shots of BG505 SOSIP N276D, and several of them developed VRC01-class responses to the CD4 binding site. Serum from these mice were tested by ELISA for the capacity to bind to two variants of BG505 MD39 that each have mutations in the VRC01 epitope that abrogate VRC01-class binding ("KO4" mutations). Thus serum binding to these two variants report on the strength of off-target, non-VRC01-class responses. This ELISA analysis shows that the sera from 4/6 mice show considerably stronger reactivity to BG505 MD39 N276D KO4 compared to BG505 MD39 GRSF4 KO4, demonstrating that a significant portion of the off-target response in these mice is directed to epitopes in the region masked by the additional GRSF4 glycans.

[0039] FIG. 18A-D depicts properties of several cleavage-independent trimers. FIG. 18A. Antigenic profile of the cleavage-independent native-like trimer BG505 CP1.2 GRSF4 compared to the cleavage-dependent native-like trimer BG505 MD39 and the non-native trimer BG505 gp120 foldon. Binding was assessed by SPR to various bnAbs and the non-nAbs 2557, 3074, 4025, F425, B6 and 17B. Trimers were analytes and IgG were captured on the sensor chip as ligand. Binding was assessed at 1 μ M trimer analyte concentration and is

displayed as the ratio of the observed response units to the response units expected for a 1 to 1 binding interaction between each gp120 on the trimer and each Fab on the IgG. Lower values correspond to less binding. FIG. 18B. Antigenic profiles of various cleavage-independent trimers compared to the cleavage-dependent trimers BG505 SOSIP, BG505 MD39 and BG505 MD39 congly, measured as in A. FIG. 18C. Expression levels of BG505 MD39 and various cleavage-independent trimers. Applicants reported in Steichen et al. that BG505 MD39 has a significantly higher expression level than BG505 SOSIP. FIG. 18D. Data illustrating that the circular permutation modification combined with the additional glycans present in MD39 CP1.2 GRSF4 serve to block undesired responses induced by soluble trimers. ELISA titrations are shown for a single VRC01 gH mice that received four immunizations : eOD-GT8 60mer, core N276D 60mer, and two shots of BG505 MD39 N276D. The mouse serum binds equally well to BG505 MD39 N276D (which has a native CD4 binding site and VRC01 epitope except for the elimination of the N276 glycan which improves VRC01-class binding) and BG505 MD39 VRC01KO4 (which has mutations in the VRC01 epitope that abrogate VRC01-class binding). Thus, the dominant serum response in this mouse is focused to off-target, non-VRC01-class epitopes on the trimer. The serum shows very little reactivity to BG505 MD39 CP1.2 GRSF4, demonstrating that the modifications in this protein serve to block the majority of off-target responses in this mouse.

[0040] FIGS. 19A-D depicts properties of the nanoparticle BG505 MD39 2JD6, an 8-mer of MD39 trimers based on the ferritin from *Pyrococcus Furiosus*. FIG. 19A. SEC-MALS trace illustrating that the purified particle in solution has the molecular weight expected of a particle composed of eight MD39 trimers each genetically fused via linkers to the 24-mer ferritin. FIG. 19B. Negative stain electron microscopy (EM) of BG505 MD39 2JD6, illustrating the existence of ordered particles. FIG. 19C. Negative stain electron microscopy (EM) of BG505 MD39 2JD6, illustrating the existence of ordered particles, at lower magnification and depicting a different sample field than in B. FIG. 19D. Antigenic profile of BG505 MD39 2JD6, assessed by SPR. BG505 MD39 2JD6 particles were captured on the sensor surface and Fabs of different antibodies were used as analytes to assess monovalent binding interactions. The SPR kinetic traces and the measured dissociation constants (K_Ds) for four bnAbs are shown on the left. SPR kinetics for two non-bnAbs are shown at right, and in those cases the binding signal was significantly weaker than for the bnAbs (note the different scale on the y-axis).

[0041] FIGS. 20A-D depicts properties of the nanoparticle BG505 MD39 E2p, a 20-mer of MD39 trimers based on the Dihydrolipoyl Transacetylase (E2p) from *Bacillus Stearothermophilus*. FIG. 20A. SEC-MALS trace illustrating that the purified particle in solution has the molecular weight expected of a particle composed of 20 MD39 trimers each genetically fused via linkers to the underlying E2p 60-mer. FIG. 20B. Negative stain electron microscopy (EM) of BG505 MD39 E2p, illustrating the existence of ordered particles. FIG. 20C. Negative stain electron microscopy (EM) of BG505 MD39 E2p, illustrating the existence of ordered particles, at lower magnification and depicting a different sample field than in B. FIG. 20D. Antigenic profile of BG505 MD39 E2p, assessed by SPR. BG505 MD39 E2p particles were captured on the sensor surface and Fabs of different antibodies were used as analytes to assess monovalent binding interactions. The SPR kinetic traces and the measured dissociation constants (K_Ds) for five bnAbs are shown on the left. SPR kinetics for two non-bnAbs are shown at right, and in those cases the binding signal was significantly weaker than for the bnAbs (note the different scale on the y-axis).

[0042] FIG. 21A-B depict flow cytometry data demonstrating cell-surface expression and bnAb-binding of membrane-bound native-like trimers. FIG. 21A. Binding of bnAbs to membrane-bound stabilized trimers using native transmembrane domains. FIG. 21B. Binding of bnAbs to membrane-bound stabilized trimers anchored to the membrane by flexible linkers and PDGFR transmembrane domains.

[0043] FIGS. 22A-F depict results of NHP immunizations with BG505 Olio6 and BG505 Olio6 CD4-KO compared to BG505 SOSIP and two other comparator trimers (BG505 SOSIP v4.1 and BG505 SOSIP v5.2). The sequence and biophysical properties of BG505 SOSIP.v4.1 were reported in De Taeye et al. Cell 2015. The sequence and biophysical properties of BG505 SOSIP.v5.2 are not known. Rhesus macaques (6 animals per immunogen group) were immunized at 3 time points: week 0, week 8, and week 24. All immunizations were administered as split doses. Each immunization consisted of 2 subcutaneous injections of 50 µg of Env trimer protein + 37.5 units (U) of Iscomatrix adjuvant (CSL) in sterile phosphate-buffered saline (PBS) diluent for a total of 100 µg of Env trimer protein + 75 U of Iscomatrix per immunization per animal. Subcutaneous immunizations were given in a volume of 0.5mL with a 1 inch, 25 gauge needle at the medial inner mid-thigh of each leg. Throughout the figure immunogen names are abbreviated as follows: BG505 SOSIP.664 (.664 or BG505 WT), BG505 SOSIP.v4.1 (v4.1), BG505 SOSIP.v5.2 (v5.2), BG505 SOSIP Olio6 (Olio6), and BG505 SOSIP Olio6 CD4-KO (Olio6

CD4-KO). All nAb titer and ELISA binding Ab data panels show geometric mean titers with geometric SD. ns, non-significant. Also see Fig. 23. FIG. 22A. BG505 V3 loop peptide binding IgG titers in BG505 SOSIP.664, BG505 SOSIP.v4.1, BG505 SOSIP.v5.2, BG505 Olio6, and BG505 Olio6 CD4-KO immunized RMs two weeks after the 3rd immunization. Olio6 and Olio6 CD4-KO both elicited significantly lower V3 binding antibody responses than BG505 SOSIP, indicating that the Olio6 and Olio6 CD4-KO retained native structure in vivo to a greater than BG505 SOSIP. Horizontal dotted line indicates limit of detection. N = 6 or 12. FIG. 22B. BG505 SOSIP binding IgG titers in BG505 SOSIP.664, BG505 SOSIP.v4.1, BG505 SOSIP.v5.2, BG505 Olio6, and BG505 Olio6 CD4-KO immunized RMs two weeks after the 3rd immunization. Olio6 and Olio6 CD4-KO both elicited significantly lower BG505 SOSIP-binding antibody responses than BG505 SOSIP. Horizontal dotted line indicates limit of detection. Week 26, N = 6 or 12. FIG. 22C. Ratio of BG505 SOSIP to BG505 V3-peptide titers (as shown panels B and C) in immunized RMs two weeks after the 3rd immunization. Olio6 and Olio6 CD4-KO both elicited responses with a significantly higher ratio than BG505 SOSIP, reflecting the strong suppression of V3 responses by Olio6 and Olio6 CD4-KO. Week 26, N = 6 or 12. FIG. 22D. Tier 1 SF162 nAb titers in BG505 SOSIP.664, BG505 SOSIP.v4.1, BG505 SOSIP.v5.2, BG505 Olio6, and BG505 Olio6 CD4-KO immunized animals two weeks after the 3rd immunization. Olio6 and Olio6 CD4-KO both elicited significantly lower SF162 tier 1 neutralization responses than BG505 SOSIP, also indicating that Olio6 and Olio6 CD4-KO retained native structure in vivo to a greater than BG505 SOSIP. Week 26, N = 6 or 12. FIG. 22E. BG505 nAb titers in BG505 SOSIP.664, BG505 SOSIP.v4.1, BG505 SOSIP.v5.2, BG505 Olio6, and BG505 Olio6 CD4-KO immunized animals two weeks after the 3rd immunization. Week 26, N = 6 or 12. FIG. 22F. Ratio of BG505 and SF162 nAb titers (as shown in panels E and F) in immunized RMs two weeks after the 3rd immunization. Olio6 and Olio6 CD4-KO both elicited responses with a significantly higher ratio than BG505 SOSIP, further indicating that Olio6 and Olio6 CD4-KO retained native structure in vivo to a greater than BG505 SOSIP. Week 26, N = 6 or 12.

[0044] FIGS. 23A-D depict additional results of NHP immunizations with BG505 Olio6 and BG505 Olio6 CD4-KO compared to BG505 SOSIP and two other comparator trimers (BG505 SOSIP v4.1 and BG505 SOSIP v5.2). The data in this figure pertain to the same RM immunization experiment described in Fig. 22. Throughout the figure immunogen names are abbreviated as follows: BG505 SOSIP.664 (.664), BG505 SOSIP.v4.1 (v4.1), BG505 SOSIP.v5.2 (v5.2), BG505 SOSIP Olio6 (Olio6), and BG505 SOSIP Olio6 CD4-KO (Olio6

CD4-KO). ns, non-significant. All nAb titer and ELISA binding Ab data panels show geometric mean titers with geometric SD. FIG. 23A. BG505 gp120 binding IgG titer 2 weeks after the 3rd immunization. Olio6 and Olio6 CD4-KO both elicited significantly lower gp120 antibody binding responses than BG505 SOSIP, further indicating that Olio6 and Olio6 CD4-KO retained native structure in vivo to a greater than BG505 SOSIP. Week 26, N = 6 or 12. FIG. 23B. Comparison of V3-peptide ELISA titers against the BG505 Olio6 V3 loop peptide (Olio6) and the BG505 SOSIP.664 V3 loop peptide (WT). V3 loop peptide binding IgG titers for both of these peptides were measured from plasma of RMs immunized with BG505 SOSIP.664, BG505 Olio6, or BG505 Olio6 CD4-KO, two weeks after the 3rd immunization. Olio6 and Olio6 CD4-KO both elicited significantly lower V3 binding antibody responses than BG505 SOSIP, regardless of whether the ELISA measurement is performed with WT or Olio6 V3 peptides. Week 26, N = 6. FIG. 23C. The ratio of BG505 nAb titer to V3-peptide binding IgG titer after the 3rd immunization. Week 26, N = 6 or 12. FIG. 23D. Tier 1 MW965 nAb titers after the 3rd immunization. Week 26, N = 6 or 12.

[0045] FIGS. 24A-B depict properties of BG505_MD39_CP1.2_GRSF7_qLoops1. FIG. 24A. SECMAALS analysis showing that the molecule is a trimer in solution. FIG. 24B. SPR analysis showing that the trimer exhibits a native-like antigenic profile. In particular the trimer binds well to the trimer-structure-specific bnAb PGT145, as well as two other bnAbs VRC01 and PGT121, but does not exhibit significant affinity for non-neutralizing Abs B6 or 4025.

DETAILED DESCRIPTION OF THE INVENTION

[0046] The invention pertains to the identification, design, synthesis and isolation of mutant trimers disclosed herein as well as nucleic acids encoding the same. The present invention also relates to homologues, derivatives and variants of the sequences of the mutant trimers and nucleic acids encoding the same, wherein it is preferred that the homologue, derivative or variant have at least 50%, at least 60%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 93%, at least 95%, at least 97%, at least 98% or at least 99% homology or identity with the sequence of the mutant trimers and nucleic acids encoding the same. It is noted that within this specification, homology to sequences of the mutant proteins and nucleic acids encoding the same refers to the homology of the homologue, derivative or variant to the binding site of the mutant proteins and nucleic acids encoding the same.

[0047] The invention still further relates to nucleic acid sequences expressing the mutant trimers disclosed herein, or homologues, variants or derivatives thereof. One of skill in the art will know, recognize and understand techniques used to create such. Additionally, one of skill in the art will be able to incorporate such a nucleic acid sequence into an appropriate vector, allowing for production of the amino acid sequence of mutant proteins and nucleic acids encoding the same or a homologue, variant or derivative thereof.

[0048] Where used herein and unless specifically indicated otherwise, the following terms are intended to have the following meanings in addition to any broader (or narrower) meanings the terms might enjoy in the art:

[0049] The term "isolated" or "non-naturally occurring" is used herein to indicate that the isolated moiety (e.g. peptide or compound) exists in a physical milieu distinct from that in which it occurs in nature. For example, the isolated peptide may be substantially isolated with respect to the complex cellular milieu in which it naturally occurs. The absolute level of purity is not critical, and those skilled in the art may readily determine appropriate levels of purity according to the use to which the peptide is to be put. The term "isolating" when used a step in a process is to be interpreted accordingly.

[0050] In many circumstances, the isolated moiety will form part of a composition (for example a more or less crude extract containing many other molecules and substances), buffer system, matrix or excipient, which may for example contain other components (including proteins, such as albumin).

[0051] In other circumstances, the isolated moiety may be purified to essential homogeneity, for example as determined by PAGE or column chromatography (for example HPLC or mass spectrometry). In preferred embodiments, the isolated peptide or nucleic acid of the invention is essentially the sole peptide or nucleic acid in a given composition.

[0052] In an advantageous embodiment, a tag may be utilized for purification or biotinylation. The tag for purification may be a his tag. In another embodiment, the tag for biotinylation may be an avi-tag. Other tags are contemplated for purification, however, purification may be accomplished without a tag. In another embodiment, antibody (such as, not limited to, a broadly neutralizing antibody) affinity columns are contemplated. In another embodiment, lectin columns are contemplated.

[0053] Native-like soluble trimers can be made by several methods that all involve stabilizing associations between envelope protein subunits. See, e.g., P. Dosenovic et al., "Immunization for HIV-1 broadly neutralizing antibodies in human Ig knockin mice," Cell,

161:1-11, 2015; J.G. Jardine et al., "Priming a broadly neutralizing antibody response to HIV-1 using a germline targeting immunogen," *Science*, doi:10.1126/science.aac5894, 2015 and R.W. Sanders et al., "HIV-1 neutralizing antibodies induced by native-like envelope trimers," *Science*, doi:10.1126/science.aac4223, 2015.

[0054] The proteins and compounds of the invention need not be isolated in the sense defined above, however.

[0055] The term "pharmaceutical composition" is used herein to define a solid or liquid composition in a form, concentration and level of purity suitable for administration to a patient (e.g. a human patient) upon which administration it may elicit the desired physiological changes. The terms "immunogenic composition" and "immunological composition" and "immunogenic or immunological composition" cover any composition that elicits an immune response against the targeted pathogen, HIV. Terms such as "vaccinal composition" and "vaccine" and "vaccine composition" cover any composition that induces a protective immune response against the targeted pathogen or which efficaciously protects against the pathogen; for instance, after administration or injection, elicits a protective immune response against the targeted pathogen or provides efficacious protection against the pathogen. Accordingly, an immunogenic or immunological composition induces an immune response, which may, but need not, be a protective immune response. An immunogenic or immunological composition may be used in the treatment of individuals infected with the pathogen, e.g., to stimulate an immune response against the pathogen, such as by stimulating antibodies against the pathogen. Thus, an immunogenic or immunological composition may be a pharmaceutical composition. Furthermore, when the text speaks of "immunogen, antigen or epitope", an immunogen may be an antigen or an epitope of an antigen. A diagnostic composition is a composition containing a compound or antibody, e.g., a labeled compound or antibody, that is used for detecting the presence in a sample, such as a biological sample, e.g., blood, semen, vaginal fluid, etc, of an antibody that binds to the compound or an immunogen, antigen or epitope that binds to the antibody; for instance, an anti-HIV antibody or an HIV immunogen, antigen or epitope.

[0056] A "conservative amino acid change" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (e.g. lysine, arginine and histidine), acidic side chains (e.g. aspartic acid and glutamic acid), non-charged amino acids or polar side chains (e.g. glycine,

asparagine, glutamine, serine, threonine, tyrosine and cysteine), non-polar side chains (e.g. alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine and tryptophan), beta-branched side chains (e.g. threonine, valine and isoleucine), and aromatic side chains (e.g. tyrosine, phenylalanine, tryptophan and histidine).

[0057] The terms “protein”, “peptide”, “polypeptide”, and “amino acid sequence” are used interchangeably herein to refer to polymers of amino acid residues of any length. The polymer may be linear or branched, it may comprise modified amino acids or amino acid analogs, and it may be interrupted by chemical moieties other than amino acids. The terms also encompass an amino acid polymer that has been modified naturally or by intervention; for example disulfide bond formation, glycosylation, lipidation, acetylation, phosphorylation, or any other manipulation or modification, such as conjugation with a labeling or bioactive component.

[0058] As used herein, the terms “antigen” or “immunogen” are used interchangeably to refer to a substance, typically a protein, which is capable of inducing an immune response in a subject. The term also refers to proteins that are immunologically active in the sense that once administered to a subject (either directly or by administering to the subject a nucleotide sequence or vector that encodes the protein) is able to evoke an immune response of the humoral and/or cellular type directed against that protein.

[0059] The term “antibody” includes intact molecules as well as fragments thereof, such as Fab, F(ab')₂, Fv and scFv which are capable of binding the epitope determinant. These antibody fragments retain some ability to selectively bind with its antigen or receptor and include, for example:

- (a) Fab, the fragment which contains a monovalent antigen-binding fragment of an antibody molecule may be produced by digestion of whole antibody with the enzyme papain to yield an intact light chain and a portion of one heavy chain;
- (b) Fab', the fragment of an antibody molecule may be obtained by treating whole antibody with pepsin, followed by reduction, to yield an intact light chain and a portion of the heavy chain; two Fab' fragments are obtained per antibody molecule;
- (c) F(ab')₂, the fragment of the antibody that may be obtained by treating whole antibody with the enzyme pepsin without subsequent reduction; F(ab')₂ is a dimer of two Fab' fragments held together by two disulfide bonds;
- (d) scFv, including a genetically engineered fragment containing the variable region of a heavy and a light chain as a fused single chain molecule.

[0060] General methods of making these fragments are known in the art. (See for example, Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, New York (1988), which is incorporated herein by reference). Fabs, Fv and scFV may also be made recombinantly, i.e. expressed as Fab, Fv or scFV rather than cleaving an intact IgG.

[0061] A “neutralizing antibody” may inhibit the entry of HIV-1 virus for example SF162 and/or JR-CSF with a neutralization index >1.5 or >2.0 . Broad and potent neutralizing antibodies may neutralize greater than about 50% of HIV-1 viruses (from diverse clades and different strains within a clade) in a neutralization assay. The inhibitory concentration of the monoclonal antibody may be less than about 25 mg/ml to neutralize about 50% of the input virus in the neutralization assay.

[0062] An “isolated antibody” or “non-naturally occurring antibody” is one that has been separated and/or recovered from a component of its natural environment. Contaminant components of its natural environment are materials that would interfere with diagnostic or therapeutic uses for the antibody, and may include enzymes, hormones, and other proteinaceous or nonproteinaceous solutes. In preferred embodiments, the antibody is purified: (1) to greater than 95% by weight of antibody as determined by the Lowry method, and most preferably more than 99% by weight; (2) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence by use of a spinning cup sequenator; or (3) to homogeneity by SDS-PAGE under reducing or non-reducing conditions using Coomassie blue or, preferably, silver stain. Isolated antibody includes the antibody *in situ* within recombinant cells since at least one component of the antibody’s natural environment will not be present. Ordinarily, however, isolated antibody will be prepared by at least one purification step.

[0063] The term “monoclonal antibody” as used herein refers to an antibody obtained from a population of substantially homogeneous antibodies, i.e., the individual antibodies which may comprise the population are identical except for possible naturally occurring mutations that may be present in minor amounts. Monoclonal antibodies are highly specific, being directed against a single antigenic site. Furthermore, in contrast to polyclonal antibody preparations that include different antibodies directed against different determinants (epitopes), each monoclonal antibody is directed against a single determinant on the antigen. In addition to their specificity, the monoclonal antibodies are advantageous in that they may be synthesized uncontaminated by other antibodies. The modifier “monoclonal” is not to be

construed as requiring production of the antibody by any particular method. For example, the monoclonal antibodies useful in the present invention may be prepared by the hybridoma methodology first described by Kohler et al., *Nature*, 256:495 (1975), or may be made using recombinant DNA methods in bacterial, eukaryotic animal or plant cells (see, *e.g.*, U.S. Pat. No. 4,816,567). The “monoclonal antibodies” may also be isolated from phage antibody libraries using the techniques described in Clackson *et al.*, *Nature*, 352:624-628 (1991) and Marks et al., *J. Mol. Biol.*, 222:581-597 (1991), for example.

[0064] An “antibody fragment” may comprise a portion of an intact antibody, preferably the antigen binding or variable region of the intact antibody. Examples of antibody fragments include Fab, Fab', F(ab')₂, scFV and Fv fragments; diabodies; linear antibodies (see U.S. Pat. No. 5,641,870; Zapata et al., *Protein Eng.* 8(10): 1057-1062 [1995]); single-chain antibody molecules; and multispecific antibodies formed from antibody fragments.

[0065] It should be understood that the proteins, including the antibodies of the invention may differ from the exact sequences illustrated and described herein. Thus, the invention contemplates deletions, additions and substitutions to the sequences shown, so long as the sequences function in accordance with the methods of the invention. In this regard, particularly preferred substitutions will generally be conservative in nature, *i.e.*, those substitutions that take place within a family of amino acids. For example, amino acids are generally divided into four families: (1) acidic--aspartate and glutamate; (2) basic--lysine, arginine, histidine; (3) non-polar-- alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan; and (4) uncharged polar--glycine, asparagine, glutamine, cysteine, serine threonine, tyrosine. Phenylalanine, tryptophan, and tyrosine are sometimes classified as aromatic amino acids. It is reasonably predictable that an isolated or non-naturally occurring replacement of leucine with isoleucine or valine, or vice versa; an aspartate with a glutamate or vice versa; a threonine with a serine or vice versa; or a similar conservative replacement of an amino acid with a structurally related amino acid, will not have a major effect on the biological activity. Proteins having substantially the same amino acid sequence as the sequences illustrated and described but possessing minor amino acid substitutions that do not substantially affect the immunogenicity of the protein are, therefore, within the scope of the invention.

[0066] As used herein the terms “nucleotide sequences” and “nucleic acid sequences” refer to deoxyribonucleic acid (DNA) or ribonucleic acid (RNA) sequences, including, without limitation, messenger RNA (mRNA), DNA/RNA hybrids, or synthetic nucleic acids.

The nucleic acid may be single-stranded, or partially or completely double-stranded (duplex). Duplex nucleic acids may be homoduplex or heteroduplex.

[0067] As used herein the term “transgene” may be used to refer to “recombinant” nucleotide sequences that may be derived from any of the nucleotide sequences encoding the proteins of the present invention. The term “recombinant” means a nucleotide sequence that has been manipulated “by man” and which does not occur in nature, or is linked to another nucleotide sequence or found in a different arrangement in nature. It is understood that manipulated “by man” means manipulated by some artificial means, including by use of machines, codon optimization, restriction enzymes, etc.

[0068] For example, in one embodiment the nucleotide sequences may be mutated such that the activity of the encoded proteins in vivo is abrogated. In another embodiment the nucleotide sequences may be codon optimized, for example the codons may be optimized for human use. In preferred embodiments the nucleotide sequences of the invention are both mutated to abrogate the normal in vivo function of the encoded proteins, and codon optimized for human use. For example, each of the sequences of the invention, such as the mutant trimers, may be altered in these ways.

[0069] As regards codon optimization, the nucleic acid molecules of the invention have a nucleotide sequence that encodes the antigens of the invention and may be designed to employ codons that are used in the genes of the subject in which the antigen is to be produced. Many viruses, including HIV and other lentiviruses, use a large number of rare codons and, by altering these codons to correspond to codons commonly used in the desired subject, enhanced expression of the antigens may be achieved. In a preferred embodiment, the codons used are “humanized” codons, i.e., the codons are those that appear frequently in highly expressed human genes (Andre et al., J. Virol. 72:1497-1503, 1998) instead of those codons that are frequently used by HIV. Such codon usage provides for efficient expression of the transgenic HIV proteins in human cells. Any suitable method of codon optimization may be used. Such methods, and the selection of such methods, are well known to those of skill in the art. In addition, there are several companies that will optimize codons of sequences, such as Geneart (geneart.com). Thus, the nucleotide sequences of the invention may readily be codon optimized.

[0070] The invention further encompasses nucleotide sequences encoding functionally and/or antigenically equivalent variants and derivatives of the antigens of the invention and functionally equivalent fragments thereof. These functionally equivalent variants, derivatives,

and fragments display the ability to retain antigenic activity. For instance, changes in a DNA sequence that do not change the encoded amino acid sequence, as well as those that result in conservative substitutions of amino acid residues, one or a few amino acid deletions or additions, and substitution of amino acid residues by amino acid analogs are those which will not significantly affect properties of the encoded polypeptide. Conservative amino acid substitutions are glycine/alanine; valine/isoleucine/leucine; asparagine/glutamine; aspartic acid/glutamic acid; serine/threonine/methionine; lysine/arginine; and phenylalanine/tyrosine/tryptophan. In one embodiment, the variants have at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98% or at least 99% homology or identity to the antigen, epitope, immunogen, peptide or polypeptide of interest.

[0071] For the purposes of the present invention, sequence identity or homology is determined by comparing the sequences when aligned so as to maximize overlap and identity while minimizing sequence gaps. In particular, sequence identity may be determined using any of a number of mathematical algorithms. A nonlimiting example of a mathematical algorithm used for comparison of two sequences is the algorithm of Karlin & Altschul, Proc. Natl. Acad. Sci. USA 1990; 87: 2264-2268, modified as in Karlin & Altschul, Proc. Natl. Acad. Sci. USA 1993;90: 5873-5877.

[0072] Another example of a mathematical algorithm used for comparison of sequences is the algorithm of Myers & Miller, CABIOS 1988;4: 11-17. Such an algorithm is incorporated into the ALIGN program (version 2.0) which is part of the GCG sequence alignment software package. When utilizing the ALIGN program for comparing amino acid sequences, a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4 may be used. Yet another useful algorithm for identifying regions of local sequence similarity and alignment is the FASTA algorithm as described in Pearson & Lipman, Proc. Natl. Acad. Sci. USA 1988; 85: 2444-2448.

[0073] Advantageous for use according to the present invention is the WU-BLAST (Washington University BLAST) version 2.0 software. WU-BLAST version 2.0 executable programs for several UNIX platforms may be downloaded from <ftp://blast.wustl.edu/blast/executables>. This program is based on WU-BLAST version 1.4, which in turn is based on the public domain NCBI-BLAST version 1.4 (Altschul & Gish, 1996, Local alignment statistics, Doolittle ed., Methods in Enzymology 266: 460-480;

Altschul et al., *Journal of Molecular Biology* 1990; 215: 403-410; Gish & States, 1993; *Nature Genetics* 3: 266-272; Karlin & Altschul, 1993; *Proc. Natl. Acad. Sci. USA* 90: 5873-5877; all of which are incorporated by reference herein).

[0074] The various recombinant nucleotide sequences and antibodies of the invention are made using standard recombinant DNA and cloning techniques. Such techniques are well known to those of skill in the art. See for example, "Molecular Cloning: A Laboratory Manual", second edition (Sambrook et al. 1989).

[0075] The nucleotide sequences of the present invention may be inserted into "vectors." The term "vector" is widely used and understood by those of skill in the art, and as used herein the term "vector" is used consistent with its meaning to those of skill in the art. For example, the term "vector" is commonly used by those skilled in the art to refer to a vehicle that allows or facilitates the transfer of nucleic acid molecules from one environment to another or that allows or facilitates the manipulation of a nucleic acid molecule.

[0076] Any vector that allows expression of the antibodies of the present invention may be used in accordance with the present invention. In certain embodiments, the antibodies of the present invention may be used in vitro (such as using cell-free expression systems) and/or in cultured cells grown in vitro in order to produce the encoded HIV- antibodies, which may then be used for various applications such as in the production of proteinaceous vaccines. For such applications, any vector that allows expression of the antibodies in vitro and/or in cultured cells may be used.

[0077] For applications where it is desired that the antibodies be expressed in vivo, for example when the transgenes of the invention are used in DNA or DNA-containing vaccines, any vector that allows for the expression of the antibodies of the present invention and is safe for use in vivo may be used. In preferred embodiments the vectors used are safe for use in humans, mammals and/or laboratory animals.

[0078] For the antibodies of the present invention to be expressed, the protein coding sequence should be "operably linked" to regulatory or nucleic acid control sequences that direct transcription and translation of the protein. As used herein, a coding sequence and a nucleic acid control sequence or promoter are said to be "operably linked" when they are covalently linked in such a way as to place the expression or transcription and/or translation of the coding sequence under the influence or control of the nucleic acid control sequence. The "nucleic acid control sequence" may be any nucleic acid element, such as, but not limited to promoters, enhancers, IRES, introns, and other elements described herein that

direct the expression of a nucleic acid sequence or coding sequence that is operably linked thereto. The term “promoter” will be used herein to refer to a group of transcriptional control modules that are clustered around the initiation site for RNA polymerase II and that when operationally linked to the protein coding sequences of the invention lead to the expression of the encoded protein. The expression of the transgenes of the present invention may be under the control of a constitutive promoter or of an inducible promoter, which initiates transcription only when exposed to some particular external stimulus, such as, without limitation, antibiotics such as tetracycline, hormones such as ecdysone, or heavy metals. The promoter may also be specific to a particular cell-type, tissue or organ. Many suitable promoters and enhancers are known in the art, and any such suitable promoter or enhancer may be used for expression of the transgenes of the invention. For example, suitable promoters and/or enhancers may be selected from the Eukaryotic Promoter Database (EPDB).

[0079] The vectors used in accordance with the present invention should typically be chosen such that they contain a suitable gene regulatory region, such as a promoter or enhancer, such that the antibodies of the invention may be expressed.

[0080] Any suitable vector may be used depending on the application. For example, plasmids, viral vectors, bacterial vectors, protozoal vectors, insect vectors, baculovirus expression vectors, yeast vectors, mammalian cell vectors, and the like, may be used. Suitable vectors may be selected by the skilled artisan taking into consideration the characteristics of the vector and the requirements for expressing the antibodies under the identified circumstances.

[0081] In an advantageous embodiment, IgG1 and Fab expression vectors may be utilized to reconstitute heavy and light chain constant regions if heavy and light chain genes of the antibodies of the present invention are cloned.

[0082] When the aim is to express the antibodies of the invention in vivo in a subject, for example in order to generate an immune response against an HIV-1 antigen and/or protective immunity against HIV-1, expression vectors that are suitable for expression on that subject, and that are safe for use in vivo, should be chosen. For example, in some embodiments it may be desired to express the antibodies of the invention in a laboratory animal, such as for pre-clinical testing of the HIV-1 immunogenic compositions and vaccines of the invention. In other embodiments, it will be desirable to express the antibodies of the invention in human subjects, such as in clinical trials and for actual clinical use of the immunogenic compositions

and vaccine of the invention. Any vectors that are suitable for such uses may be employed, and it is well within the capabilities of the skilled artisan to select a suitable vector. In some embodiments it may be preferred that the vectors used for these in vivo applications are attenuated to vector from amplifying in the subject. For example, if plasmid vectors are used, preferably they will lack an origin of replication that functions in the subject so as to enhance safety for in vivo use in the subject. If viral vectors are used, preferably they are attenuated or replication-defective in the subject, again, so as to enhance safety for in vivo use in the subject.

[0083] In preferred embodiments of the present invention viral vectors are used. Viral expression vectors are well known to those skilled in the art and include, for example, viruses such as adenoviruses, adeno-associated viruses (AAV), alphaviruses, herpesviruses, retroviruses and poxviruses, including avipox viruses, attenuated poxviruses, vaccinia viruses, and particularly, the modified vaccinia Ankara virus (MVA; ATCC Accession No. VR-1566). Such viruses, when used as expression vectors are innately non-pathogenic in the selected subjects such as humans or have been modified to render them non-pathogenic in the selected subjects. For example, replication-defective adenoviruses and alphaviruses are well known and may be used as gene delivery vectors.

[0084] The nucleotide sequences and vectors of the invention may be delivered to cells, for example if the aim is to express the HIV-1 antigens in cells in order to produce and isolate the expressed proteins, such as from cells grown in culture. For expressing the antibodies in cells any suitable transfection, transformation, or gene delivery methods may be used. Such methods are well known by those skilled in the art, and one of skill in the art would readily be able to select a suitable method depending on the nature of the nucleotide sequences, vectors, and cell types used. For example, transfection, transformation, microinjection, infection, electroporation, lipofection, or liposome-mediated delivery could be used. Expression of the antibodies may be carried out in any suitable type of host cells, such as bacterial cells, yeast, insect cells, and mammalian cells. The antibodies of the invention may also be expressed using including in vitro transcription/translation systems. All of such methods are well known by those skilled in the art, and one of skill in the art would readily be able to select a suitable method depending on the nature of the nucleotide sequences, vectors, and cell types used.

[0085] A synthetic mutant trimer may be chemically synthesized in whole or part using techniques that are well-known in the art (see, e.g., Kochendoerfer, G. G., 2001). Additionally, homologs and derivatives of the polypeptide may be also be synthesized.

[0086] Alternatively, methods which are well known to those skilled in the art may be used to construct expression vectors containing nucleic acid molecules that encode the polypeptide or homologs or derivatives thereof under appropriate transcriptional/translational control signals, for expression. These methods include in vitro recombinant DNA techniques, synthetic techniques and in vivo recombination/genetic recombination. See, for example, the techniques described in Maniatis et al., 1989.

[0087] The HIV envelope protein (Env) is the target of broadly neutralizing antibodies (bnAbs) in natural infection. Env is a membrane protein composed of a trimer of gp120 and gp41 subunits that contains a high degree of sequence diversity and a surface that is shielded by N-linked glycans. The bnAbs that target Env often have unusual features such as a long complementarity-determining region (CDR) H3, high levels of somatic hypermutation (SHM), and insertions and deletions (INDELS). Furthermore, most of the bnAbs recognize complex epitopes that are typically non-linear and have both protein and glycan components.

[0088] The most common epitope of bnAbs in HIV infected individuals is a high mannose glycan patch at the base of the variable loop V3 that includes a glycan linked to N332 (Landais et al. 2016 PLoS Pathog. 12, e1005369). PGT121 and its somatic relatives are an exceptionally potent family of bnAbs that target this epitope and PGT121 has been shown to protect macaques in SHIV challenge studies (Walker et al. 2011 Nature. 477, 466-470, Moldt et al. 2012 Proc Natl Acad Sci. 109, 18921-18925). The elicitation of high and sustained titers of PGT121-like antibodies by vaccination would therefore have a reasonable likelihood of providing protection against HIV in humans.

[0089] HIV Env proteins show no detectable affinity for predicted germline precursors of PGT121, suggesting that activation of appropriate precursors is a barrier to PGT121-like bnAb induction that could be addressed by germline-targeting immunogen design. In this view, vaccine induction of PGT121-like bnAbs might be achieved by a germline-targeting prime followed by boosts with progressively more native-like Env, similar to what has been proposed for elicitation of VRC01-class bnAbs (Jardine, Julien, Menis et al. 2013 Science. 340, 711-716; McGuire et al 2013 J Exp Med. 210, 655-663; Jardine, Ota, Sok et al. 2015 Science. 349, 156-161). BG505 SOSIP.664 was the first soluble native-like Env trimer (Sanders et al. 2013 PloS Pathog. 9, e1003618). In parallel with Applicants' germline-targeting effort a goal was to improve the expression, stability and antigenic profile of BG505 SOSIP.664 in order to have an enhanced trimer platform for germline targeting and boosting. Using a lentivirus-based method for displaying libraries of immunogens on the surface of

mammalian cells, and guided by the known structure of BG505 SOSIP.664 (Julien et al. 2013 Science. 342, 1477-1483; Lyumkis et al. 2013 Science. 342, 1484-1490; Pancera et al. 2014 Nature. 514, 455-461) Applicants have engineered a series of soluble native-like trimers with improved yield, thermostability and antigenic profile, which have progressively increasing affinity for putative PGT121 germline precursors and intermediately mutated antibodies.

[0090] Applicants have demonstrated that structure-guided mammalian cell surface display can be used to engineer trimers containing native-like glycans. Native-like trimers have been developed that bind to predicted PGT121 germline precursors and intermediately mutated antibodies. BG505 SOSIP trimers were engineered with improved yield, thermostability and antigenic profile. Tests of priming and boosting strategies are currently underway in PGT121-GL knock-in mice.

[0091] Applicants claim sequences of different types of immunogen sequences. The sequences provided below are exemplary examples, the stabilizing mutations, modifications, (such as, but not limited to, cleavage-independent modifications), and/or a membrane anchoring strategy (such as, but not limited to, linker plus platelet-derived growth factor receptor (PDGFR)) described herein are applicable to any HIV strain or clade, such as but not limited to, those described below.

[0092] As used herein, at least three separate zoonotic transmissions resulted in the formation of three distinct HIV-1 groups: M (main), O (outlier), and N (non-M/non-O).

[0093] About 90% of HIV-1 infections are classified as group M and these are distributed worldwide. Group O infections are endemic to several west central African countries and represent 1 to 5% of all HIV-1 infection in those areas. Group N has only been identified in a small number of individuals in Cameroon.

[0094] Within the HIV-M group, there is a further division into at least ten subtypes or clades (groups of genetically related virus). Historically, the distribution of subtypes followed the geographic patterns listed below.

[0095] Clade or Subtype A: Central and East Africa as well as East European countries that were formerly part of the Soviet Union.

[0096] Clade or Subtype B: West and Central Europe, the Americas, Australia, South America, and several southeast Asian countries (Thailand, and Japan), as well as northern Africa and the Middle East.

[0097] Clade or Subtype C: Sub-Saharan Africa, India, and Brazil.

[0098] Clade or Subtype D: North Africa and the Middle East.

[0099] Clade or Subtype F: South and southeast Asia.

[00100] Clade or Subtype G: West and Central Africa.

[00101] Clade or Subtypes H, J, and K: Africa and the Middle East.

[00102] Additionally, different subtypes can combine genetic material to form a hybrid virus, known as a “circulating recombinant form” (CRFs), of which at least twenty have been identified (see, e.g., 2.Buonarguro L Human Immunodeficiency Virus Type 1 Subtype distribution in the worldwide epidemic: pathogenetic and therapeutic implications. J Virol 81(19):10209-19, 2007).

[00103] The present invention encompasses the stabilizing mutations, modifications, (such as, but not limited to, cleavage-independent modifications), and/or a membrane anchoring strategy (such as, but not limited to, linker plus platelet-derived growth factor receptor (PDGFR)) described herein to all groups and clades of HIV.

[00104] Types I and II are gp120 molecules (I) and gp140 trimer molecules (II) with mutations discovered to improve binding to germline-reverted and/or less-mutated versions of PGT121. The sequences in I and II can be employed in sequential immunization schemes to attempt to elicit PGT121-class bnAbs against HIV.

[00105] Type III are gp140 trimer molecules with stabilizing mutations to increase expression level and/or increase thermal melting temperature and/or improve antigenic profile, where a favorable antigenic profile means better affinity for broadly neutralizing antibodies and no or very weak affinity for non-neutralizing antibodies.

[00106] Type IV are combinations of mutations from II and III: these are gp140 trimers that contain both stabilizing mutations and germline-targeting mutations. In type IV Applicants have listed only a few important combinations, but the present invention encompasses all possible combinations of the mutations from II and III.

[00107] Type V are trimers with modified surfaces or of different strains than BG505, that can be employed in strategic boosting regimens.

[00108] Type VI are additional trimer modifications that add extra functionality and that can be combined with types II, III, IV or V.

[00109] Type VII are examples of native-like trimers from other HIV strains that have been stabilized by MD39 and Olio6 mutations, demonstrating the general applicability of the MD39 and Olio6 stabilizing mutations. FIG. 16A provides experimental data supporting these trimers as native-like.

[00110] Type VIII are variants of BG505 MD39 that do not require cleavage by furin. We refer to these as "cleavage-independent" trimers.

[00111] Type IX are glycan masked trimers in which N-linked glycosylation sites have been added to cover the bottom and sides of the soluble trimer.

[00112] Type X are native-like trimers with variable loops V1, V2b and V4 modified to both minimize their lengths and maximize the number of glycosylation sites contained within them.

[00113] Type XI are BG505 MD39-based, single-component, self-assembling nanoparticles.

[00114] Type XII are BG505 MD39-based, membrane-bound native like trimers.

[00115] Amino acid and nucleic acid sequences are listed below. In the amino acid sequences, mutations relative to a parent construct are generally indicated in bold.

[00116] I: gp120s with PGT121-class germline-targeting mutations

[00117] BG505-gp120-L111A-2_T135A_T139I_mC (BG505-gp120 3mut)

[00118] VWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLENVTEEFNMWKN
NMVEQMHTDIISAWDQSLKPCVKLTPLCVTLQCTNVANNIIDDMRGELKNCSFNMTTEL RDK
KQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIHYCAPAG
FAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAE EVMIRSENITNNAKNI
LVQFNTPVQINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAH CNVSKATWNETLGKVVK
QLRKHFNGNTIIRFANSSGGDLEVTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQGSNSTGS
NDSITLPCRIKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGG
GDMRDNRSELYKVKVVKIEP

[00119] GTGTGGAAAGACGCCGAGACTACTCTGTTCTGTGCCAGCGACGCCAAAGCATACG
AAACTGAAAAGCATAATGTGTGGGCTACTCACGCCTGCGTGCCACAGACCCAAATCCCCAG
GAAATCCACCTGGAGAATGTCACTGAGGAATTCAACATGTGGAAGAACAATATGGTGGAGCA
GATGCATACCGACATCATTTTCAGCCTGGGATCAGAGCCTGAAGCCATGCGTGAAACTGACTC
CCCTGTGCGTCACCCCTGCAGTGTACCAACGTGGCCAAACAATATCATCGACGATATGCGGGGC
GAACTGAAGAATTGTTCTTTCAACATGACCACAGAGCTGCGGGACAAGAAACAGAAAGTGTA
CAGTCTGTTTTATAGACTGGATGTGGTCCAGATCAATGAAAACCAGGGGAATCGATCCAACA
ATTCTAACAAGGAGTACCGGCTGATCAATTGCAACACTAGCGCCATTACCCAGGCTTGTCCT
AAAGTGTCTTTGAACCTATCCCAATTCATTATTGCGCCCCTGCTGGCTTCGCCATCCTGAA
GTGTAAAGATAAGAAGTTCAACGGCACCGGGCCCTGCCCTTCTGTGAGTACAGTCCAGTGTA
CTCACGGGATTAAGCCTGTGGTCAGTACACAGCTGCTGCTGAATGGATCACTGGCTGAGGAA
GAAGTGATGATCCGATCCGAGAACATTACTAACAATGCAAAGAATATCCTGGTGCAGTTCAA
CACCCCTGTCCAGATTAATTGCACTCGCCCAAACAATAACACCCGAAAAGCATCAGAATTG
GACCAGGCCAGGCATTTTACGCCACCGGGGACATCATTTGGAGATATCAGACAGGCACACTGT

AATGTGTCCAAGGCCACCTGGAACGAAACACTGGGAAAGGTGGTCAAACAGCTGAGAAAACA
TTTCGGCAATAAACTATCATTAGGTTTGCTAATAGCTCCGGCGGGGACCTGGAAGTGACTA
CCCACAGCTTCAACTGCGGAGGCGAGTTCTTTTACTGTAAACACATCCGGCCTGTTTAATTCT
ACATGGATCAGTAACACTTCAGTGCAGGGCTCAAATAGCACCGGGAGCAACGATTCCATCAC
ACTGCCTTGCCGCATTAAGCAGATCATTAAATATGTGGCAGCGAATTGGACAGGCTATGTATG
CACCCCTATCCAGGGCGTGATTAGATGTGTCTCTAATATCACCGGGCTGATTCTGACACGC
GACGGGGGATCTACAAACAGTACAACTGAGACTTTCAGGCCAGGCGGGGGAGACATGAGGGA
TAACTGGCGCAGCGAACTGTACAAGTATAAAGTGGTCAAATCGAGCCC

[00120] BG505-gp120-L111A-2_5mut_mC (BG505-gp120 5mut)

[00121] VWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLENVTEEFNMWKN
NMVEQMHTDIIISAWDQSLKPCVKLTPLCVTLQCTNYT**PNLT**NDMRGELKNCSFNMTTEL RDK
KQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIHYCAPAG
FAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEEEVMIRSENITNNAKNI
LVQFNTFPVQINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNVSKATWNETLGKVVK
QLRKHFNGNTIIRFANSSGGDLEVTTHSFNCGEFFYCNTSGLFNSTWISNTSVQGSNSTGS
NDSITLPCRICKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGG
GDMRDNRSELYKYKVVKIEP

[00122] GTGTGGAAGACGCCGAGACTACTCTGTTCTGTGCCAGCGACGCCAAAGCATAACG
AAACTGAAAAGCATAATGTGTGGGCTACTCACGCCTGCGTGCCACAGACCCAAATCCCAG
GAAATCCACCTGGAGAATGTCACTGAGGAATTCAACATGTGGAAGAACAATATGGTGGAGCA
GATGCATACCGACATCATTTACGCCTGGGATCAGAGCCTGAAGCCATGCGTGAAACTGACTC
CCCTGTGCGTCACCCTGCAGTGTACCAACTACACACCCAATCTGACCAACGATATGCGGGGC
GAACTGAAGAATTGTTCTTTCAACATGACCACAGAGCTGCGGGACAAGAAACAGAAAGTGTA
CAGTCTGTTTTATAGACTGGATGTGGTCCAGATCAATGAAAACCAGGGGAATCGATCCAACA
ATTCTAACAAGGAGTACCGGCTGATCAATTGCAACACTAGCGCCATTACCCAGGCTTGTCCC
AAAGTGTCTTTGAACCTATCCCAATTCATTATTGCGCCCCTGCTGGCTTCGCCATCCTGAA
GTGTAAAGATAAGAAGTTCAACGGCACCGGGGCCCTGCCCTTCTGTGAGTACAGTCCAGTGTA
CTCACGGGATTAAGCCTGTGGTCAGTACACAGCTGCTGCTGAATGGATCACTGGCTGAGGAA
GAAGTGATGATCCGATCCGAGAACATTACTAACAATGCAAAGAATATCCTGGTGCAGTTCAA
CACCCCTGTCCAGATTAATTGCACTCGCCCAAACAATAACACCCGAAAAGCATCAGAATTG
GACCAGGCCAGGCATTTTACGCCACCGGGGACATCATTGGAGATATCAGACAGGCACACTGT
AATGTGTCCAAGGCCACCTGGAACGAAACACTGGGAAAGGTGGTCAAACAGCTGAGAAAACA
TTTCGGCAATAAACTATCATTAGGTTTGCTAATAGCTCCGGCGGGGACCTGGAAGTGACTA
CCCACAGCTTCAACTGCGGAGGCGAGTTCTTTTACTGTAAACACATCCGGCCTGTTTAATTCT
ACATGGATCAGTAACACTTCAGTGCAGGGCTCAAATAGCACCGGGAGCAACGATTCCATCAC
ACTGCCTTGCCGCATTAAGCAGATCATTAAATATGTGGCAGCGAATTGGACAGGCTATGTATG
CACCCCTATCCAGGGCGTGATTAGATGTGTCTCTAATATCACCGGGCTGATTCTGACACGC

GACGGGGGATCTACAAACAGTACAAGTACTGAGACTTTTCAGGCCAGGCGGGGGAGACATGAGGGA
TAACTGGCGCAGCGAACTGTACAAGTATAAAGTGGTCAAAATCGAGCCC

[00123] BG505-gp120-L111A-2_7mut_mC (BG505-gp120 7mut or BG505 gp120 7MUT)

[00124] VWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQEIHLENVTEEFNMWKN
NMVEQMHTDIIISAWDQSLKPCVKLTPLCVTLQCTNY**APNL**INDMRGELKNCSFNMTTEL RDK
KQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIHYCAPAG
FAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAE E EVMIRSENITNNAKNI
LVQFNTPVQINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAH CNVSKATWNETLGKVVK
QLRKHFNGNTIIRFANSSGGDLEVTTHSFNCGGEFFYCN TSGLFNSTWISNTSVQGSNSTGS
NDSITLPCR KQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGG
GDMRDNRSELYKYKVVKIEP

[00125] GTGTGGAAAGACGCCGAGACTACTCTGTTCTGTGCCAGCGACGCCAAAGCATACG
AACTGAAAAGCATAATGTGTGGGCTACTCACGCCTGCGTGCCACAGACCCAAATCCCCAG
GAAATCCACCTGGAGAATGTCACTGAGGAATTCAACATGTGGAAGAACAATATGGTGGAGCA
GATGCATACCGACATCATTTTCAGCCTGGGATCAGAGCCTGAAGCCATGCGTGAAACTGACTC
CCCTGTGCGTCACCTGTCAGTGTACCAACTACGCCCCCAATCTGATCAACGATATGCGGGGC
GAACTGAAGAATTGTTCTTTCAACATGACCACAGAGCTGCGGGACAAGAAACAGAAAGTGTA
CAGTCTGT TTTATAGACTGGATGTGGTCCAGATCAATGAAAACCAGGGGAATCGATCCAACA
ATTCTAACAAGGAGTACCGGCTGATCAATTGCAACACTAGCGCCATTACCCAGGCTTGTCCT
AAAGTGTCTTTGAACCTATCCCAATTCAATTATTGCGCCCCTGCTGGCTTCGCCATCCTGAA
GTGTAAAGATAAGAAGTTCAACGGCACCGGGGCCCTGCCCTTCTGTGAGTACAGTCCAGTGTA
CTCACGGGATTAAGCCTGTGGTCAGTACACAGCTGCTGCTGAATGGATCACTGGCTGAGGAA
GAAGTGATGATCCGATCCGAGAACATTACTAACAATGCAAAGAATATCCTGGTGCAGTTCAA
CACCCCTGTCCAGATTAATTGCACTCGCCCAACAATAACACCCGGAAAAGCATCAGAATTG
GACCAGGCCAGGCATTTTACGCCACCGGGGACATCATTTGGAGATATCAGACAGGCACACTGT
AATGTGTCCAAGGCCACCTGGAACGAAACACTGGGAAAGGTGGTCAAACAGCTGAGAAAACA
TTTCGGCAATAACACTATCATTAGGTTTGCTAATAGCTCCGGCGGGGACCTGGAAGTGACTA
CCCACAGCTTCAACTGCGGAGGCGAGTTCTTTTACTGTAACACATCCGGCCTGTTTAATTCT
ACATGGATCAGTAACACTTCAGTGCAGGGCTCAAATAGCACCGGGAGCAACGATTCCATCAC
ACTGCCTTGCCGCATTAAGCAGATCATTAATATGTGGCAGCGAATTGGACAGGCTATGTATG
CACCCCTATCCAGGGCGTGATTAGATGTGTCTCTAATATCACCGGGCTGATTCTGACACGC
GACGGGGGATCTACAAACAGTACAAGTACTGAGACTTTTCAGGCCAGGCGGGGGAGACATGAGGGA
TAACTGGCGCAGCGAACTGTACAAGTATAAAGTGGTCAAAATCGAGCCC

[00126] BG505-gp120-L111A-2_10mut_mC (BG505-gp120 10mut or BG505 gp120 10MUT)

[00127] VWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQEIHLENVTEEFNMWKN
NMVEQMHTDIIISAWDQSLKPCVKLTPLCVTLQCTNY**APFL**INDMRGELKNCSFNMTTEL RDK
KQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIHYCAPAG

FAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLNGSLAEEVMI RSENITNNAKNI
 LVQFNTPVQINCTRPNNNTRKSIRIGPGQAFYA**FGDI**IGDIR**MA**HNCNVSKATWNETLGKVVK
 QLRKHFGNNTIIRFANSSGGDLEVTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQGSNSTGS
 NDSITLPCRIKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGG
 GDMRDNRSELYKYKVVKIEP

[00128] GTGTGGAAAGACGCCGAGACTACTCTGTTCTGTGCCAGCGACGCCAAAGCATACG
 AAAGTGAAGCATAATGTGTGGGCTACTCACGCCTGCGTGCCACAGACCCAAATCCCCAG
 GAAATCCACCTGGAGAATGTCACTGAGGAATTCAACATGTGGAAGAACAATATGGTGGAGCA
 GATGCATACCGACATCATTTACGCCTGGGATCAGAGCCTGAAGCCATGCGTGAAACTGACTC
 CCCTGTGCGTCACCCCTGCAGTGACCAACTACGCCCCCTTCCTGATCAACGATATGCGGGGC
 GAACTGAAGAATTGTTCTTTCAACATGACCACAGAGCTGCGGGACAAGAAACAGAAAGTGTA
 CAGTCTGTTTTATAGACTGGATGTGGTCCAGATCAATGAAAACCAGGGGAATCGATCCAACA
 ATTCTAACAAGGAGTACCGGTGATCAATTGCAACACTAGCGCCATTACCCAGGCTTGTCCC
 AAAGTGTCTTTGAACCTATCCCAATTCATTATTGCGCCCCTGCTGGCTTCGCCATCCTGAA
 GTGTAAAGATAAGAAGTTCAACGGCACCCGGGCCCTGCCCTTCTGTGAGTACAGTCCAGTGTA
 CTCACGGGATTAAGCCTGTGGTCAGTACACAGCTGCTGCTGAATGGATCACTGGCTGAGGAA
 GAAGTGATGATCCGATCCGAGAACATTACTAACAATGCAAAGAATATCCTGGTGCAGTTCAA
 CACCCCTGTCCAGATTAATTGCACTCGCCCCAAACAATAACACCCGGAAAAGCATCAGAATTG
 GACCAGGCCAGGCATTTTACGCCTTCGGGGACATCATTTGGAGATATCAGAATGGCACACTGT
 AATGTGTCCAAGGCCACCTGGAACGAAACACTGGGAAAGGTGGTCAAACAGCTGAGAAAACA
 TTTCGGCAATAAACAATCATTAGGTTTGCTAATAGCTCCGGCGGGGACCTGGAAGTGACTA
 CCCACAGCTTCAACTGCGGAGGCGAGTTCTTTTACTGTAAACACATCCGGCCTGTTTAATTCT
 ACATGGATCAGTAACACTTCAGTGCAGGGCTCAAATAGCACCGGGAGCAACGATTCCATCAC
 ACTGCCTTGCCGCATTAAGCAGATCATTAATATGTGGCAGCGAATTGGACAGGCTATGTATG
 CACCCCTATCCAGGGCGTGATTAGATGTGTCTCTAATATCACCGGGCTGATTCTGACACGC
 GACGGGGGATCTACAAACAGTACAACCTGAGACTTTACAGGCCAGGCGGGGGAGACATGAGGGA
 TAACTGGCGCAGCGAACTGTACAAGTATAAAGTGGTCAAATCGAGCCC

[00129] BG505-gp120-L111A-2_10mut2A_mC

[00130] VWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQEIHLNVTEEFNMWKN
 NMVEQMHTDIIISAWDQSLKPCVKLTPLCVTLQCTNY**APFLIN**MRGELKNCSFNMTTTEL RDK
 KQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIHYCAPAG
 FAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLNGSLAEEVMI RSENITNNAKNI
 LVQFNTPVQINCTRPNNNTRKSIRIGPGQAFYA**FGDI**IGDIR**MA**HNCNVSKATWNETLGKVVK
 QLRKHFGNNTIIRFANSSGGDLEVTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQGSNSTGS
 NDSITLPCRIKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGG
 GDMRDNRSELYKYKVVKIEP

[00131] GTGTGGAAAGACGCCGAGACTACTCTGTTCTGTGCCAGCGACGCCAAAGCATACG
 AAAGTGAAGCATAATGTGTGGGCTACTCACGCCTGCGTGCCACAGACCCAAATCCCCAG

GAAATCCACCTGGAGAATGTCACTGAGGAATTCAACATGTGGAAGAACAATATGGTGGAGCA
 GATGCATACCGACATCATTTTCAGCCTGGGATCAGAGCCTGAAGCCATGCGTGAAACTGACTC
 CCCTGTGCGTCACCTGTCAGTGTACCAACTACGCCCCCTTCCTGATCAACAACATGCGGGGC
 GAACTGAAGAATTGTTCTTTCAACATGACCACAGAGCTGCGGGACAAGAAACAGAAAGTGTA
 CAGTCTGTTTTATAGACTGGATGTGGTCCAGATCAATGAAAACCAGGGGAATCGATCCAACA
 ATTCTAACAAGGAGTACCGGCTGATCAATTGCAACACTAGCGCCATTACCCAGGCTTGTCCTC
 AAAGTGTCTTTGAACCTATCCCAATTCATTATTGCGCCCCCTGCTGGCTTCGCCATCCTGAA
 GTGTAAAGATAAGAAGTTCAACGGCACCGGGCCCTGCCCTTCTGTGAGTACAGTCCAGTGTA
 CTCACGGGATTAAGCCTGTGGTCAGTACACAGCTGCTGCTGAATGGATCACTGGCTGAGGAA
 GAAGTGATGATCCGATCCGAGAACATTACTAACAATGCAAAGAATATCCTGGTGCAGTTCAA
 CCCCCCTGTCCAGATTAATTGCACTCGCCCAAACAATAACACCCGGAAAAGCATCAGAATTG
 GACCAGGCCAGGCATTTTACGCCTTCGGGGACATCATTTGGAGATATCAGAAATGGCACACTGT
 AATGTGTCCAAGGCCACCTGGAACGAAACACTGGGAAAGGTGGTCAAACAGCTGAGAAAACA
 TTTCGGCAATAACACTATCATTAGGTTTGCTAATAGCTCCGGCGGGGACCTGGAAGTGACTA
 CCCACAGCTTCAACTGCGGAGGCGAGTTCTTTTACTGTAACACATCCGGCCTGTTTAATTCT
 ACATGGATCAGTAACACTTCAGTGCAGGGCTCAAATAGCACCGGGAGCAACGATTCCATCAC
 ACTGCCTTGCCGCATTAAGCAGATCATTAATATGTGGCAGCGAATTGGACAGGCTATGTATG
 CACCCCTATCCAGGGCGTGATTAGATGTGTCTCTAATATCACCGGGCTGATTCTGACACGC
 GACGGGGGATCTACAAACAGTACAACCTGAGACTTTCAGGCCAGGCGGGGGAGACATGAGGGA
 TAACTGGCGCAGCGAAGTGTACAAGTATAAAGTGGTCAAATCGAGCCC

[00132] BG505-gp120-L111A-2_11mut2A_mC

[00133] VWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLNVTEEFNMWKN
 NMVEQMHTDIISAWDQSLKPCVKLTPLCVTLQCTNYAPNLLSNMRGELKNCSFNMTELRDK
 KQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIHYCAPAG
 FAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEEVMI RSENITNNAKNI
 LVQFNTPVQINCTRPNNNTRKSIRIGPGQAFYAFGDIIGDIRMAHCNVSKATWNETLGKVVK
 QLRKHFGNNTIIRFANSSGGDLEVTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQGSNSTGS
 NDSIVLPCRIKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGG
 GDMRDNRSELYKYKVVKIEP

[00134] GTGTGGAAGACGCCGAGACTACTCTGTTCTGTGCCAGCGACGCCAAAGCATACG
 AAAGTGAAGCATAATGTGTGGGCTACTCACGCCTGCGTGCCACAGACCCAAATCCCAG
 GAAATCCACCTGGAGAATGTCACTGAGGAATTCAACATGTGGAAGAACAATATGGTGGAGCA
 GATGCATACCGACATCATTTTCAGCCTGGGATCAGAGCCTGAAGCCATGCGTGAAACTGACTC
 CCCTGTGCGTCACCTGTCAGTGTACCAACTACGCCCCAACCTGCTGAGCAACATGCGGGGC
 GAACTGAAGAATTGTTCTTTCAACATGACCACAGAGCTGCGGGACAAGAAACAGAAAGTGTA
 CAGTCTGTTTTATAGACTGGATGTGGTCCAGATCAATGAAAACCAGGGGAATCGATCCAACA
 ATTCTAACAAGGAGTACCGGCTGATCAATTGCAACACTAGCGCCATTACCCAGGCTTGTCCTC
 AAAGTGTCTTTGAACCTATCCCAATTCATTATTGCGCCCCCTGCTGGCTTCGCCATCCTGAA

GTGTAAAGATAAGAAAGTTCAACGGCACCGGGCCCTGCCCTTCTGTGAGTACAGTCCAGTGTA
 CTCACGGGATTAAGCCTGTGGTCAGTACACAGCTGCTGCTGAATGGATCACTGGCTGAGGAA
 GAAGTGATGATCCGATCCGAGAACATTACTAACAATGCAAAGAATATCCTGGTGAGTTCAA
 CCCCCCTGTCCAGATTAATTGCACTCGCCCAAACAATAACACCCGGAAAAGCATCAGAATTG
 GACCAGGCCAGGCATTTTACGCCTTCGGGGACATCATTGGAGATATCAGAATGGCACACTGT
 AATGTGTCCAAGGCCACCTGGAACGAAACACTGGGAAAGGTGGTCAAACAGCTGAGAAAACA
 TTTCGGCAATAACACTATCATTAGGTTTGCTAATAGCTCCGGCGGGGACCTGGAAGTACTA
 CCCACAGCTTCAACTGCGGAGGCGAGTTCTTTTACTGTAACACATCCGGCCTGTTTAATTCT
 ACATGGATCAGTAACACTTCAGTGCAGGGCTCAAATAGCACCGGGAGCAACGATTCCATCGT
 GCTGCCTTGCCGCATTAAGCAGATCATTAAATATGTGGCAGCGAATTGGACAGGCTATGTATG
 CCCCCCTATCCAGGGCGTGATTAGATGTGTCTCTAATATCACCGGGCTGATTCTGACACGC
 GACGGGGGATCTACAAACAGTACAACAGTACTTTCAGGCCAGGCGGGGGAGACATGAGGGA
 TAACTGGCGCAGCGAACTGTACAAGTATAAAGTGGTCAAATCGAGCCC

[00135] II: Trimers with PGT121-class germline-targeting mutations

[00136] BG505_SOSIP.D664_JS_3mut_mC (SOSIP-3MUT)

[00137] AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIH
 LENVT EEFNMWKNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLQCTNVANNIIDD MRGELK
 NCSFNMTTEL RDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVS
 FEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEFEV
 M IRSENITNNAKNILVQFNTPVQINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNVS
 KATWNETLGKVVKQLRKHFGNNTIIRFANSSGGDLEVTHSFNCGGEFFYCNTSGLFNSTWI
 SNTSVQGSNSTGSNDSITLPCR IKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGG
 STNSTTETFRPGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRAVGIGAVF
 LGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQQHLLKLTWVGIKQLQARVL
 AVERYLRDQQLLGIWCGSKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
 GLLEESQNQQEKNEQDLLALD

[00138] GCTGAAAACCTGTGGGTCACTGTCTACTACGGCGTGCCTGTCTGGAAGGATGCTG
 AAATACTCTGTTCTGTGCCTCTGATGCTAAAGCCTACGAAACCGAGAAGCACAATGTGTGG
 GCCACACATGCTTGCGTCCCTACTGACCCAAACCCCGAGGAAATCCACCTGGAGAATGTGAC
 AGAGGAATTCAACATGTGGA AAAACAATATGGTGGAGCAGATGCATACTGACATCATTTCCC
 TGTGGGATCAGTCTCTGAAGCCTTGCGTGAACTGACACCACTGTGCGTCACTCTGCAGTGT
 ACTAACGTGGCCAACAATATCATCGACGATATGCGGGGCGAACTGAAGAATTGTTCTTTCAA
 CATGACCACAGAGCTGCGGGACAAGAAACAGAAAGTGACAGTCTGTTTTATAGACTGGATG
 TGGTCCAGATCAATGAAAACAGGGGAATCGAAGTAACAATTCAAACAAGGAGTACCGGCTG
 ATCAATTGCAACACAAGTGCAATTACTCAGGCCTGTCCAAAAGTGTCATTTGAACCTATCCC
 AATTCAATTATTGCGCTCCCGCAGGCTTCGCCATCCTGAAGTGTAAGATAAGAAGTTCAACG
 GCACCGGGCCATGCCCTTCAGTGAGCACCGTCCAGTGACACACGGAATTAAGCCAGTGGTC
 TCCACACAGCTGCTGCTGAATGGCTCTCTGGCTGAGGAAGAAGTGATGATCAGGAGCGAGAA

CATTACCAACAATGCAAAGAATATCCTGGTGCAGTTCAACACACCCCGTCCAGATTAATTGCA
 CCCGCCCTAACAATAACACACGGAAATCTATCAGAATTGGACCCGGCCAGGCATTTTATGCC
 ACCGGCGACATCATTGGGGATATCAGACAGGCACACTGTAATGTGAGCAAGGCCACTTGGA
 CGAGACCCTGGGGAAGGTGGTCAAACAGCTGAGAAAGCATTTCGGAAACAACACTATCATCA
 GGTTTGCCAATAGCTCCGGCGGGGACCTGGAAGTGACTACCCACAGCTTCAACTGCGGAGGC
 GAGTTCCTTTACTGTAACACAAGTGGCCTGTTTAATTCAACCTGGATCAGCAACACATCCGT
 GCAGGGATCCAATTCTACAGGCTCTAACGATAGTATCACTCTGCCCTGCCGCATTAAGCAGA
 TCATTAATATGTGGCAGCGAATCGGGCAGGCTATGTATGCACCCCTATCCAGGGAGTGATT
 AGATGTGTCAGCAATATCACTGGCCTGATTCTGACCCGCGACGGGGGATCAACTAACAGCAC
 AACTGAGACCTTCCGGCCTGGAGGAGGAGACATGAGGGATAACTGGCGCTCCGAACTGTACA
 AGTATAAAGTGGTCAAGATCGAGCCACTGGGAGTGGCACCAACCAGATGCAAACGAAGAGTG
 GTCGGAAGGCGACGACGAAGAAGGGCAGTGGGAATTGGAGCTGTCTTCCTGGGCTTTCTGGG
 GGCCGCTGGATCTACAATGGGGGACAGCCAGTATGACCCTGACAGTCCAGGCTCGAAATCTGC
 TGTCAGGAATCGTGCAGCAGCAGAGCAACCTGCTGAGAGCCCCTGAAGCTCAGCAGCATCTG
 CTGAAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCTAGGGTGCTGGCAGTCGAGCGCTA
 CCTGCGAGATCAGCAGCTGCTGGGCATCTGGGGGTGTTCCGGAAAGCTGATTTGCTGTACCA
 ATGTGCCATGGAACCTAGTTGGAGCAATCGCAACCTGTCCGAAATCTGGGACAATATGACT
 TGGCTGCAGTGGGATAAGGAGATTAGCAACTACCCAGATCATCTACGGCCTGCTGGAAGA
 GTCCCAGAATCAGCAGGAGAAGAACGAGCAGGACCTGCTGGCCCTGGAT

[00139] BG505_SOSIP.D664_JS_5mut_mC (SOSIP-5MUT)

[00140] AENLWVTVYYGVVPWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIH
 LENVT EEFNMWKNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLQCTNYTPNLTNDMRGELK
 NCSFNMTTEL RDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVS
 FEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAE E EVM
 IRSENITNNAKNILVQFNTPVQINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNVS
 KATWNETLGKVVQQLRKHFGNNTIIRFANSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWI
 SNTSVQGSNSTGSNDSITLPCRICKQIINMWQRIGQAMYAPPIQGVI RCVSNITGLILTRDGG
 STNSTTETFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRAVGIGAVF
 LGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQQHLLKLTWVGIKQLQARVL
 AVERYLRDQQLLGIWCGSKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
 GLLEESQNQQEKNEQDLLALD

[00141] GCTGAAAACCTGTGGGTCACTGTCTACTACGGCGTGCCTGTCTGGAAGGATGCTG
 AAATACTCTGTTCTGTGCCTCTGATGCTAAAGCCTACGAAACCGAGAAGCACAATGTGTGG
 GCCACACATGCTTGCGTCCCTACTGACCCAAACCCCGAGAAATCCACCTGGAGAATGTGAC
 AGAGGAATTCAACATGTGGAAAAACAATATGGTGGAGCAGATGCATACTGACATCATTTCCC
 TGTGGGATCAGTCTCTGAAGCCTTGCGTGAACTGACACCACTGTGCGTCACTCTGCAGTGT
 ACTAACTACACCCCAATCTGACCAACGATATGCGGGCGAACTGAAGAATTGTTCTTTCAA
 CATGACCACAGAGCTGCGGGACAAGAAACAGAAAGTGACAGTCTGTTTTATAGACTGGATG

TGGTCCAGATCAATGAAAACCAGGGGAATCGAAGTAACAATTCAAACAAGGAGTACCGGCTG
ATCAATTGCAACACAAGTGCAATTACTCAGGCCTGTCCAAAAGTGTCAATTTGAACCTATCCC
AATTCATTATTGCGCTCCCGCAGGCTTCGCCATCCTGAAGTGTAAGATAAGAAGTTCAACG
GCACCGGGCCATGCCCTTCAGTGAGCACCGTCCAGTGACACACGGAATTAAGCCAGTGGTC
TCCACACAGCTGCTGCTGAATGGCTCTCTGGCTGAGGAAGAAGTGATGATCAGGAGCGAGAA
CATTACCAACAATGCAAAGAATATCCTGGTGCAGTTCAACACACCCGTCCAGATTAATTGCA
CCCGCCCTAACAATAACACACGGAAATCTATCAGAATTGGACCCGGCCAGGCATTTTATGCC
ACCGGCGACATCATTGGGGATATCAGACAGGCACACTGTAATGTGAGCAAGGCCACTTGAA
CGAGACCCTGGGGAAGGTGGTCAAACAGCTGAGAAAGCATTTTCGAAACAACACTATCATCA
GGTTTGCCAATAGCTCCGGCGGGGACCTGGAAGTGACTACCCACAGCTTCAACTGCGGAGGC
GAGTTCTTTTACTGTAACACAAGTGGCCTGTTTAATTCAACCTGGATCAGCAACACATCCGT
GCAGGGATCCAATTCTACAGGCTCTAACGATAGTATCACTCTGCCCTGCCGCATTAAGCAGA
TCATTAATATGTGGCAGCGAATCGGGCAGGCTATGTATGCACCCCTATCCAGGGAGTGATT
AGATGTGTCAGCAATATCACTGGCCTGATTCTGACCCGCGACGGGGGATCAACTAACAGCAC
AACTGAGACCTTCCGGCCTGGAGGAGGAGACATGAGGGATAACTGGCGCTCCGAACTGTACA
AGTATAAAGTGGTCAAGATCGAGCCACTGGGAGTGGCACCAACCAGATGCAAACGAAGAGTG
GTCGGAAGGCGACGACGAAGAAGGGCAGTGGGAATTGGAGCTGTCTTCCTGGGCTTTCTGGG
GGCCGCTGGATCTACAATGGGGGCAGCCAGTATGACCCTGACAGTCCAGGCTCGAAATCTGC
TGTCAGGAATCGTGCAGCAGCAGAGCAACCTGCTGAGAGCCCCTGAAGCTCAGCAGCATCTG
CTGAAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCTAGGGTGCTGGCAGTCGAGCGCTA
CCTGCGAGATCAGCAGCTGCTGGGCATCTGGGGGTGTTCCGGAAAGCTGATTTGCTGTACCA
ATGTGCCATGGAACCTAGTTGGAGCAATCGCAACCTGTCCGAAATCTGGGACAATATGACT
TGGCTGCAGTGGGATAAGGAGATTAGCAACTACCCAGATCATCTACGGCCTGCTGGAAGA
GTCCCAGAATCAGCAGGAGAAGAACGAGCAGGACCTGCTGGCCCTGGAT

[00142] BG505_SOSIP.D664_JS_7mut_mC (SOSIP-7MUT)

[00143] AENLWVTVYYGVVPVKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIH
LENVTEEFNMWKNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLQCTNYAPNLLINDMRGELK
NCSFNMTELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVS
FEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEFEVM
IRSENITNNAKNILVQFNTPVQINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNVS
KATWNETLGKVVKQLRKHFGNNTIIRFANSSGGDLEVTHSFNCGGEFFYCNTSGLFNSTWI
SNTSVQGSNSTGSNDSITLPCRICKIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGG
STNSTTETFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVF
LGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQHLLKLTWVGIKQLQARVL
AVERYLRDQQLLGIWCGSKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
GLLEESQNQQEKNEQDLLALD

[00144] GCTGAAAACCTGTGGGTCACTGTCTACTACGGCGTGCCTGTCTGGAAGGATGCTG
AACTACTCTGTTCTGTGCCTCTGATGCTAAAGCCTACGAAACCGAGAAGCACAAATGTGTGG

GCCACACATGCTTGCGTCCCTACTGACCCAAACCCCCAGGAAATCCACCTGGAGAATGTGAC
 AGAGGAATTCAACATGTGGAAAAACAATATGGTGGAGCAGATGCATACTGACATCATTTCCC
 TGTGGGATCAGTCTCTGAAGCCTTGCGTGAACTGACACCACTGTGCGTCACTCTGCAGTGT
 ACTAACTACGCCCCCAATCTGATCAACGATATGCGGGGCGAACTGAAGAATTGTTCTTTCAA
 CATGACCACAGAGCTGCGGGACAAGAAACAGAAAGTGTACAGTCTGTTTTATAGACTGGATG
 TGGTCCAGATCAATGAAAACCAGGGGAATCGAAGTAACAATTCAAACAAGGAGTACCGGCTG
 ATCAATTGCAACACAAGTGCAATTACTCAGGCCTGTCCAAAAGTGTCAATTTGAACCTATCCC
 AATTCAATTATTGCGCTCCCGCAGGCTTCGCCATCCTGAAGTGTAAGATAAGAAGTTCAACG
 GCACCGGGCCATGCCCTTCAGTGAGCACCGTCCAGTGACACACGGAATTAAGCCAGTGGTC
 TCCACACAGCTGCTGCTGAATGGCTCTCTGGCTGAGGAAGAAGTGATGATCAGGAGCGAGAA
 CATTACCAACAATGCAAAGAATATCCTGGTGCAGTTCAACACACCCGTCCAGATTAATTGCA
 CCCGCCCTAACAATAACACACGGAAATCTATCAGAATTGGACCCGGCCAGGCATTTTATGCC
 ACCGGCGACATCATTGGGGATATCAGACAGGCACACTGTAATGTGAGCAAGGCCACTTGAA
 CGAGACCCTGGGGAAGGTGGTCAAACAGCTGAGAAAGCATTTTCGGAAACAACACTATCATCA
 GGTTTGCCAATAGCTCCGGCGGGGACCTGGAAGTGACTACCCACAGCTTCAACTGCGGAGGC
 GAGTTCTTTTACTGTAACACAAGTGGCCTGTTTAATTCAACCTGGATCAGCAACACATCCGT
 GCAGGGATCCAATTCTACAGGCTCTAACGATAGTATCACTCTGCCCTGCCGCATTAAGCAGA
 TCATTAATATGTGGCAGCGAATCGGGCAGGCTATGTATGCACCCCTATCCAGGGAGTGATT
 AGATGTGTCAGCAATATCACTGGCCTGATTCTGACCCGCGACGGGGGATCAACTAACAGCAC
 AACTGAGACCTTCCGGCCTGGAGGAGGAGACATGAGGGATAACTGGCGCTCCGAACTGTACA
 AGTATAAAGTGGTCAAGATCGAGCCACTGGGAGTGGCACCAACCAGATGCAAACGAAGAGTG
 GTCGGAAGGCGACGACGAAGAAGGGCAGTGGGAATTGGAGCTGTCTTCCTGGGCTTTCTGGG
 GGCCGCTGGATCTACAATGGGGGAGCCAGTATGACCCTGACAGTCCAGGCTCGAAATCTGC
 TGTGAGGAATCGTGAGCAGCAGAGCAACCTGCTGAGAGCCCCTGAAGCTCAGCAGCATCTG
 CTGAAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCTAGGGTGCTGGCAGTCGAGCGCTA
 CCTGCGAGATCAGCAGCTGCTGGGCATCTGGGGGTGTTCCGGAAAGCTGATTTGCTGTACCA
 ATGTGCCATGGAACCTAGTTGGAGCAATCGCAACCTGTCCGAAATCTGGGACAATATGACT
 TGGCTGCAGTGGGATAAGGAGATTAGCAACTACCCAGATCATCTACGGCCTGCTGGAAGA
 GTCCCAGAATCAGCAGGAGAAGAACGAGCAGGACCTGCTGGCCCTGGAT

[00145] BG505_SOSIP.D664_JS_10mut_mC (SOSIP-10MUT)

[00146] AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQEIH
 LENVTEEFNMWKNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLQCTNY**APFLIND**MRGELK
 NCSFNMTELRLDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVS
 FEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAE~~EE~~VM
 IRSENITNNAKNILVQFNTPVQINCTRPNNNTRKSIRIGPGQAFYA**FGDI**IGDIR**MAH**CNVS
 KATWNETLGKVVKQLRKHFGNNTIIRFANSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWI
 SNTSVQGSNSTGSNDSITLPCRICKIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGG
 STNSTTETFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVF

LGFLGAAGSTMGAASMTLTVQARNLLSGIVQQSNLLRAPEAQQHLLKLTWVGIKQLQARVL
AVERYLRDQQLLGIWCGSKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
GLLEESQNQQEKNEQDLLALD

[00147] GCTGAAAACCTGTGGGTCACTGTCTACTACGGCGTGCCTGTCTGGAAGGATGCTG
AACTACTCTGTTCTGTGCCTCTGATGCTAAAGCCTACGAAACCGAGAAGCACAATGTGTGG
GCCACACATGCTTGCGTCCCTACTGACCCAAACCCCCAGGAAATCCACCTGGAGAATGTGAC
AGAGGAATTCAACATGTGGAAAAACAATATGGTGGAGCAGATGCATACTGACATCATTTCCC
TGTGGGATCAGTCTCTGAAGCCTTGCGTGAACTGACACCACTGTGCGTCACTCTGCAGTGT
ACTAACTACGCCCCCTTTCTGATCAACGATATGCGGGCGAACTGAAGAATTGTTCTTTCAA
CATGACCACAGAGCTGCGGGACAAGAAACAGAAAGTGACAGTCTGTTTTATAGACTGGATG
TGGTCCAGATCAATGAAAACAGGGGAATCGAAGTAACAATTCAAACAAGGAGTACCGGCTG
ATCAATTGCAACACAAGTGCAATTACTCAGGCCTGTCCAAAAGTGTCAATTTGAACCTATCCC
AATTCATTATTGCGCTCCCGCAGGCTTCGCCATCCTGAAGTGTAAGATAAGAAGTTCAACG
GCACCGGGCCATGCCCTTCAGTGAGCACCGTCCAGTGACACACGGAATTAAGCCAGTGGTC
TCCACACAGCTGCTGCTGAATGGCTCTCTGGCTGAGGAAGAAGTGATGATCAGGAGCGAGAA
CATTACCAACAATGCAAAGAATATCCTGGTGCAGTTCAACACACCCGTCCAGATTAATTGCA
CCCGCCCTAACAATAACACACGGAAATCTATCAGAATTGGACCCGGCCAGGCATTTTATGCC
TTTGGCGACATCATTGGGGATATCAGAATGGCACACTGTAATGTGAGCAAGGCCACTTGAA
CGAGACCCTGGGGAAGGTGGTCAAACAGCTGAGAAAGCATTTCGGAAACAACACTATCATCA
GGTTTGCCAATAGCTCCGGCGGGGACCTGGAAGTGACTACCCACAGCTTCAACTGCGGAGGC
GAGTTCTTTTACTGTAACACAAGTGGCCTGTTTAATTCAACCTGGATCAGCAACACATCCGT
GCAGGGATCCAATTCTACAGGCTCTAACGATAGTATCACTCTGCCCTGCCGCATTAAGCAGA
TCATTAATATGTGGCAGCGAATCGGGCAGGCTATGTATGCACCCCTATCCAGGGAGTGATT
AGATGTGTCAGCAATATCACTGGCCTGATTCTGACCCGCGACGGGGGATCAACTAACAGCAC
AACTGAGACCTTCCGGCCTGGAGGAGGAGACATGAGGGATAACTGGCGCTCCGAACTGTACA
AGTATAAAGTGGTCAAGATCGAGCCACTGGGAGTGGCACCAACCAGATGCAAACGAAGAGTG
GTCGGAAGGCGACGACGAAGAAGGGCAGTGGGAATTGGAGCTGTCTTCCTGGGCTTTCTGGG
GGCCGCTGGATCTACAATGGGGGAGCCAGTATGACCCTGACAGTCCAGGCTCGAAATCTGC
TGTCAGGAATCGTGCAGCAGCAGAGCAACCTGCTGAGAGCCCCTGAAGCTCAGCAGCATCTG
CTGAAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCTAGGGTGCTGGCAGTCGAGCGCTA
CCTGCGAGATCAGCAGCTGCTGGGCATCTGGGGGTGTTCCGGAAAGCTGATTTGCTGTACCA
ATGTGCCATGGAACCTAGTTGGAGCAATCGCAACCTGTCCGAAATCTGGGACAATATGACT
TGGCTGCAGTGGGATAAGGAGATTAGCAACTACACCCAGATCATCTACGGCCTGCTGGAAGA
GTCCCAGAATCAGCAGGAGAAGAACGAGCAGGACCTGCTGGCCCTGGAT

[00148] BG505_SOSIP.D664_JS_10mut2A_mC

[00149] AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIH
LENVTEEFNMWKNNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLQCTNYAPFLINNMRGELK
NCSFNMTTELKQKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVS

FEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEEEVM
IRSENITNNAKNILVQFNTPVQINCTRPNNNTRKSIRIGPGQAFYA**F**GDIIGDIR**MA**HNCNVS
KATWNETLGKVVKQLRKHFGNNTIIRFANSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWI
SNTSVQGSNSTGSNDSITLPCRICKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGG
STNSTTETFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVF
LGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQHLLKLTWVGIKQLQARVL
AVERYLRDQQLLGIWCGSKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
GLLEESQNQQEKNEQDLLALD

[00150] GCTGAAAACCTGTGGGTCACTGTCTACTACGGCGTGCCTGTCTGGAAGGATGCTG
AACTACTCTGTTCTGTGCCTCTGATGCTAAAGCCTACGAAACCGAGAAGCACAATGTGTGG
GCCACACATGCTTGCGTCCCTACTGACCCAAACCCCCAGGAAATCCACCTGGAGAATGTGAC
AGAGGAATTCAACATGTGGAAAAACAATATGGTGGAGCAGATGCATACTGACATCATTTCCC
TGTGGGATCAGTCTCTGAAGCCTTGCGTGAACTGACACCACTGTGCGTCACTCTGCAGTGT
ACTAACTACGCCCCCTTTCTGATCAACAACATGCGGGGCGAACTGAAGAATTGTTCTTTCAA
CATGACCACAGAGCTGCGGGACAAGAAACAGAAAGTGACAGTCTGTTTTATAGACTGGATG
TGGTCCAGATCAATGAAAACAGGGGAATCGAAGTAACAATTCAAACAAGGAGTACCGGCTG
ATCAATTGCAACACAAGTGCAATTACTCAGGCCTGTCCAAAAGTGTCAATTTGAACCTATCCC
AATTCATTATTGCGCTCCCGCAGGCTTCGCCATCCTGAAGTGTAAGATAAGAAGTTCAACG
GCACCGGGCCATGCCCTTCAGTGAGCACCGTCCAGTGACACACGGAATTAAGCCAGTGGTC
TCCACACAGCTGCTGCTGAATGGCTCTCTGGCTGAGGAAGAAGTGATGATCAGGAGCGAGAA
CATTACCAACAATGCAAAGAATATCCTGGTGCAGTTCAACACACCCGTCCAGATTAATTGCA
CCCGCCCTAACAATAACACACGGAATCTATCAGAATTGGACCCGGCCAGGCATTTTATGCC
TTTGGCGACATCATTGGGGATATCAGAATGGCACACTGTAATGTGAGCAAGGCCACTTGGAA
CGAGACCTTGGGGAAGGTGGTCAAACAGCTGAGAAAGCATTTCGGAAACAACACTATCATCA
GGTTTGCCAATAGCTCCGGCGGGGACCTGGAAGTGACTACCCACAGCTTCAACTGCGGAGGC
GAGTTCTTTTACTGTAACACAAGTGGCCTGTTTAATTCAACCTGGATCAGCAACACATCCGT
GCAGGGATCCAATTCTACAGGCTCTAACGATAGTATCACTCTGCCCTGCCGCATTAAGCAGA
TCATTAATATGTGGCAGCGAATCGGGCAGGCTATGTATGCACCCCTATCCAGGGAGTGATT
AGATGTGTCAGCAATATCACTGGCCTGATTCTGACCCGCGACGGGGGATCAACTAACAGCAC
AACTGAGACCTTCCGGCCTGGAGGAGGAGACATGAGGGATAACTGGCGCTCCGAACTGTACA
AGTATAAAGTGGTCAAGATCGAGCCACTGGGAGTGGCACCAACCAGATGCAAACGAAGAGTG
GTCGGAAGGCGACGACGAAGAAGGGCAGTGGGAATTGGAGCTGTCTTCCTGGGCTTTCTGGG
GGCCGCTGGATCTACAATGGGGGCAGCCAGTATGACCCTGACAGTCCAGGCTCGAAATCTGC
TGTCAGGAATCGTGACGAGCAGAGCAACCTGCTGAGAGCCCCTGAAGCTCAGCAGCATCTG
CTGAAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCTAGGGTGCTGGCAGTGCAGCGCTA
CCTGCGAGATCAGCAGCTGCTGGGCATCTGGGGGTGTTCCGGAAAGCTGATTTGCTGTACCA
ATGTGCCATGGAACCTAGTTGGAGCAATCGCAACCTGTCCGAAATCTGGGACAATATGACT

TGGCTGCAGTGGGATAAGGAGATTAGCAACTACACCCAGATCATCTACGGCCTGCTGGAAGA
GTCCCAGAATCAGCAGGAGAAGAACGAGCAGGACCTGCTGGCCCTGGAT

[00151] BG505_SOSIP.D664_JS_11mut2A_mC

[00152] AENLWVTVYYGVVPWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIH
LENVTEEFNMWKNNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLQCTNYAPNLLSNMRGELK
NCSFNMTTEL RDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVS
FEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEFEV
IRSENITNNAKNILVQFNTPVQINCTRPNNNTRKSIRIGPGQAFYA**F**GDIIGDIR**M**AHCNVS
KATWNETLGKVVKQLRKHFGNNTIIRFANSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWI
SNTSVQGSNSTGSNDSIVLPCRICKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGG
STNSTTETFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVF
LGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQHLLKLTWVGIKQLQARVL
AVERYLRDQQLLGIWCGSKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
GLLEESQNQQEKNEQDLLALD

[00153] GCTGAAAACCTGTGGGTCACTGTCTACTACGGCGTGCCTGTCTGGAAGGATGCTG
AACTACTCTGTTCTGTGCCTCTGATGCTAAAGCCTACGAAACCGAGAAGCACAATGTGTGG
GCCACACATGCTTGCGTCCCTACTGACCCAAACCCCCAGGAAATCCACCTGGAGAATGTGAC
AGAGGAATTCAACATGTGGAAAAACAATATGGTGGAGCAGATGCATACTGACATCATTTCCC
TGTGGGATCAGTCTCTGAAGCCTTGCGTGAACTGACACCACTGTGCGTCACTCTGCAGTGT
ACTAACTACGCCCCAAACCTGCTGTCCAATATGCGGGCGAACTGAAGAATTGTTCTTTCAA
CATGACCACAGAGCTGCGGGACAAGAAACAGAAAGTGACAGTCTGTTTTATAGACTGGATG
TGGTCCAGATCAATGAAAACAGGGGAATCGAAGTAAACAATTCAAACAAGGAGTACCGGCTG
ATCAATTGCAACACAAGTGCAATTACTCAGGCCTGTCCAAAAGTGTCATTTGAACCTATCCC
AATTCAATTATTGCGCTCCCGCAGGCTTCGCCATCCTGAAGTGTAAGATAAGAAGTTCAACG
GCACCGGGCCATGCCCTTCAGTGAGCACCGTCCAGTGACACACGGAATTAAGCCAGTGGTC
TCCACACAGCTGCTGCTGAATGGCTCTCTGGCTGAGGAAGAAGTGATGATCAGGAGCGAGAA
CATTACCAACAATGCAAAGAATATCCTGGTGCAGTTCAACACACCCGTCAGATTAATTGCA
CCCGCCCTAACAATAACACACGGAAATCTATCAGAATTGGACCCGGCCAGGCATTTTATGCC
TTTGGCGACATCATTGGGGATATCAGAATGGCACACTGTAATGTGAGCAAGGCCACTTGGA
CGAGACCCTGGGGAAGGTGGTCAAACAGCTGAGAAAGCATTTTCGGAAACAACACTATCATCA
GGTTTGCCAATAGCTCCGGCGGGGACCTGGAAGTGACTACCCACAGCTTCAACTGCGGAGGC
GAGTTCTTTTACTGTAACACAAGTGGCCTGTTTAATTCAACCTGGATCAGCAACACATCCGT
GCAGGGATCCAATTCTACAGGCTCTAACGATAGTATCGTGCTGCCCTGCCGCATTAAGCAGA
TCATTAATATGTGGCAGCGAATCGGGCAGGCTATGTATGCACCCCTATCCAGGGAGTGATT
AGATGTGTCAGCAATATCACTGGCCTGATTCTGACCCGCGACGGGGGATCAACTAACAGCAC
AACTGAGACCTTCCGGCCTGGAGGAGGAGACATGAGGGATAACTGGCGCTCCGAAGTGTACA
AGTATAAAGTGGTCAAGATCGAGCCACTGGGAGTGGCACCAACCAGATGCAAACGAAGAGTG
GTCGGAAGGCGACGACGAAGAAGGGCAGTGGGAATTGGAGCTGTCTTCCTGGGCTTTCTGGG

GGCCGCTGGATCTACAATGGGGGAGCCAGTATGACCCTGACAGTCCAGGCTCGAAATCTGC
 TGTCAGGAATCGTGCAGCAGCAGAGCAACCTGCTGAGAGCCCCTGAAGCTCAGCAGCATCTG
 CTGAAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCTAGGGTGTGGCAGTCGAGCGCTA
 CCTGCGAGATCAGCAGCTGCTGGGCATCTGGGGGTGTTCCGAAAGCTGATTTGCTGTACCA
 ATGTGCCATGGAACCTAGTTGGAGCAATCGCAACCTGTCCGAAATCTGGGACAATATGACT
 TGGCTGCAGTGGGATAAGGAGATTAGCAACTACCCAGATCATCTACGGCCTGCTGGAAGA
 GTCCCAGAATCAGCAGGAGAAGAACGAGCAGGACCTGCTGGCCCTGGAT

[00154] III: Stabilized trimers with improved expression, thermal stability, and/or antigenic profile for mature bnAbs

[00155] BG505_SOSIP_D664_MD39_mC (BG505 SOSIP-MD39 or MD39)

[00156] AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHL
 LENVTEEFNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELK
 NCSFNMTELRDCKQKVYSLFYRLDVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVS
 FEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEVEVI
 IRSENITNNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNVS
 KATWNETLGKVVQQLRKHFGNNTIIRFAQSSGGDLEVTTHSFNCGGEFFYCNTSGLFNSTWI
 SNTSVQGSNSTGSNDSITLPCRICKIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGG
 STNSTTETFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVS
 LGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARVL
 AVEHYLRDQQLLGIWCGSKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
 GLLEESQNNQEKNEQDLLALD

[00157] GCCGAAAATCTGTGGGTGACTGTCTACTATGGCGTGCCTGTCTGGAAGGACGCCG
 AGACCACACTGTTCTGTGCTTCCGATGCTAAGGCATACGAAACCGAGAAACACAACGTGTGG
 GCAACCCATGCCTGCGTCCCAACAGACCCAAACCCCCAGGAAATCCACCTGGAGAATGTGAC
 CGAGGAGTTCAACATGTGGAAGAACAATATGGTGGAAACAGATGCATGAGGACATCATTTCCC
 TGTGGGATCAGTCTCTGAAGCCTTGCGTGAAACTGACCCCACTGTGCGTGACACTGCAGTGT
 ACAAACGTCACTAACAATATCACCGACGATATGCGGGGCGAACTGAAGAATTGTTCTTTCAA
 CATGACTACCGAGCTGAGGGACAAGAAACAGAAAGTGTACAGTCTGTTTTATCGCCTGGATG
 TGGTCCAGATCAATGAAAACAGGGGAATCGAAGTAACAATTCAAACAAGGAGTACCGGCTG
 ATTAATTGCAACACCACTGCCATCACACAGGCTTGTCCAAAAGTGTCAATTCGAGCCTATCCC
 AATTCATTATTGCGCCCCCGCTGGCTTCGCCATCCTGAAGTGTAAGATAAGAAGTTCAACG
 GCACCGGGCCATGCCCTTCAGTGAGCACTGTCCAGTGTACCCACGGAATTAAGCCTGTGGTC
 TCCACACAGCTGCTGCTGAATGGCTCTCTGGCTGAGGAAGAAGTGATCATTAGGTCTGAGAA
 CATCACTAACAATGCAAAGAATATTCTGGTGCAGCTGAACACCCCGTCCAGATCAATTGCA
 CTCGCCCTAACAATAACACCGTGAAATCTATCCGAATTGGACCCGGCCAGGCTTTTTATTAC
 ACCGGCGACATCATTGGCGACATCAGACAGGCACACTGCAATGTGAGCAAGGCCACATGGAA
 CGAGACTCTGGGGAAGGTGGTCAAACAGCTGCGCAAACATTTCCGAAATAACACAATCATTC
 GATTTGCACAGAGCAGCGGAGGGGACCTGGAAGTGACAACTCACAGCTTCAATTGCGGAGGC

GAGTTCCTTTTACTGTAACACTAGTGGCCTGTTTAATTCAACTTGGATCAGCAACACCTCCGT
 GCAGGGCAGCAACAGCACCGGCTCTAACGATAGTATCACACTGCCATGTCTGGATTAAGCAGA
 TCATTAACATGTGGCAGAGAATCGGGCAGGCCATGTATGCACCCCCTATCCAGGGAGTGATT
 CGATGCGTGAGCAATATCACAGGCCTGATTCTGACTAGAGACGGGGGATCAACAAACAGCAC
 CACAGAGACTTTCCGGCCCCGGCGGAGGAGACATGCGAGATAACTGGAGATCCGAAGTGTACA
 AGTATAAAGTGGTCAAGATCGAGCCACTGGGCGTGGCTCCCACCAGATGCAAACGAAGAGTG
 GTCGGGAGGCGCCGACGGAGAAGGGCTGTGGGGATTGGAGCAGTCAGCCTGGGCTTTCTGGG
 GGCCGCTGGATCTACAATGGGGGCGAGCCAGTATGACTCTGACCGTGCAGGCCAGGAATCTGC
 TGTCAGGAATCGTGCAGCAGCAGAGCAACCTGCTGAGGGCTCCCGAACCCAGCAGCACCTG
 CTGAAGGACACCCACTGGGGCATCAAGCAGCTGCAGGCAAGAGTGCTGGCCGTCGAGCATT
 CCTGAGGGATCAGCAGCTGCTGGGCATCTGGGGATGCAGCGGAAAGCTGATTTGCTGTACAA
 ATGTGCCTTGGAACCTCTAGTTGGAGCAATCGCAACCTGTCCGAAATCTGGGACAATATGACA
 TGGCTGCAGTGGGATAAGGAGATTAGCAACTACACTCAGATCATCTACGGCCTGCTGGAAGA
 GTCCCAGAATCAGCAGGAGAAGAACGAGCAGGACCTGCTGGCCCTGGAT

[00158] BG505_SOSIP_D664_MD16_mC stabilizes the V3 loop

[00159] AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIH
 LENVT EEFNMWKNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELK
 NCSFNMTTEL RDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVS
 FEPIPIHYCAPAGWAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAE E EVM
 IRSENITNNAKNILVQFNTPVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNVS
 KATWNETLGKVVQQLRKHFGNNTIIRFANSSGGDLEVTHSFNCGGEFFYCNTSGLFNSTWI
 SNTSVQGSNSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGG
 STNSTTETFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVF
 LGFLGAAGSTMGAASMTLT VQARNLLSGIVQQQSNLLRAPEAQQHLLKDTVWGIKQLQARVL
 AVERYLRDQQLLGIWCGSKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
 GLLEESQNQQEKNEQDLLALD

[00160] GCTGAAAACCTGTGGGTCACTGTCTACTACGGCGTGCCTGTCTGGAAGGATGCTG
 AAATACTCTGTTCTGTGCCTCTGATGCTAAAGCCTACGAAACCGAGAAGCACAAATGTGTGG
 GCCACACATGCTTGCGTCCCTACTGACCCAAACCCCGAGAAATCCACCTGGAGAATGTGAC
 AGAGGAATTCAACATGTGGAAAAACAATATGGTGGAGCAGATGCATACTGACATCATTTCCC
 TGTGGGATCAGTCTCTGAAGCCTTGCGTGAAACTGACACCACTGTGCGTCACTCTGCAGTGT
 ACTAACGTGACCAACAATATCACCGACGATATGCGGGGCGAACTGAAGAATTGTTCTTTCAA
 CATGACCACAGAGCTGCGGGACAAGAAACAGAAAGTGACAGTCTGTTTTATAGACTGGATG
 TGGTCCAGATCAATGAAAACAGGGGAATCGAAGTAACAATTCAAACAAGGAGTACCGGCTG
 ATCAATTGCAACACAAGTGCAATTACTCAGGCCTGTCCAAAAGTGTCATTTGAACCTATCCC
 AATTCAATTATTGCGCTCCCGCAGGCTGGGCCATCCTGAAGTGTAAGATAAGAAGTTCAACG
 GCACCGGGCCATGCCCTTCAGTGAGCACCGTCCAGTGACACACGGAATTAAGCCAGTGGTC
 TCCACACAGCTGCTGCTGAATGGCTCTCTGGCTGAGGAAGAAGTGATGATCAGGAGCGAGAA

CATTACCAACAATGCAAAGAATATCCTGGTGCAGTTCAACACACCCCGTCCAGATTAATTGCA
 CCCGCCCTAACAATAACACAGTGAAATCTATCAGAATTGGACCCGGCCAGGCATTTTATTAC
 ACCGGCGACATCATTGGGGATATCAGACAGGCACACTGTAATGTGAGCAAGGCCACTTGGA
 CGAGACCCTGGGGAAGGTGGTCAAACAGCTGAGAAAGCATTTCGGAAACAACACTATCATCA
 GGTTTGCCAATAGCTCCGGCGGGGACCTGGAAGTGACTACCCACAGCTTCAACTGCGGAGGC
 GAGTTCTTTTACTGTAACACAAGTGGCCTGTTTAATTCAACCTGGATCAGCAACACATCCGT
 GCAGGGATCCAATTCTACAGGCTCTAACGATAGTATCACTCTGCCCTGCCGCATTAAGCAGA
 TCATTAATATGTGGCAGCGAATCGGGCAGGCTATGTATGCACCCCTATCCAGGGAGTGATT
 AGATGTGTCAGCAATATCACTGGCCTGATTCTGACCCGCGACGGGGGATCAACTAACAGCAC
 AACTGAGACCTTCCGGCCTGGAGGAGGAGACATGAGGGATAACTGGCGCTCCGAACTGTACA
 AGTATAAAGTGGTCAAGATCGAGCCACTGGGAGTGGCACCAACCAGATGCAAACGAAGAGTG
 GTCGGAAGGCGACGACGAAGAAGGGCAGTGGGAATTGGAGCTGTCTTCCTGGGCTTTCTGGG
 GGCCGCTGGATCTACAATGGGGGACAGCCAGTATGACCCTGACAGTCCAGGCTCGAAATCTGC
 TGTCAGGAATCGTGCAGCAGCAGAGCAACCTGCTGAGAGCCCCTGAAGCTCAGCAGCATCTG
 CTGAAGGACACCGTGTGGGGCATCAAGCAGCTGCAGGCTAGGGTGCTGGCAGTCGAGCGCTA
 CCTGCGAGATCAGCAGCTGCTGGGCATCTGGGGGTGTTCCGGAAAGCTGATTTGCTGTACCA
 ATGTGCCATGGAACCTAGTTGGAGCAATCGCAACCTGTCCGAAATCTGGGACAATATGACT
 TGGCTGCAGTGGGATAAGGAGATTAGCAACTACCCAGATCATCTACGGCCTGCTGGAAGA
 GTCCCAGAATCAGCAGGAGAAGAACGAGCAGGACCTGCTGGCCCTGGAT

[00161] BG505_SOSIP_D664_split1_1_mC increases melting temp 10C

[00162] AENLWVTVYYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHECVPTDPNPQEIH
 LENVTEEFNMWKNMVEQMHTDI IELWDQSLKPCVKLTPLCVTLQCTNVNNTDDMRGELK
 NCSFNMTTEL RDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVS
 FEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAE E EVM
 IRSENITNNAKNILVQFNTPVQINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNVS
 KATWNETLGKVVQQLRKHFGNNTIIRFANSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWI
 SNTSVQGSNSTGSNDSITLPCRICKQIINMWQRIGQAMYAPPIQGVI RCVSNITGLILTRDGG
 STNSTTETFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRAVGIGAVF
 LGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQHLLKLTWVGIKQLQARVL
 AVERYLRDQQLLGIWCGSKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
 GLLEESQNQQEKNEQDLLALD

[00163] GCTGAAAACCTGTGGGTCACTGTCTACTACGGCGTGCCTGTCTGGAAGGATGCTG
 AAATACTCTGTTCTGTGCCTCTGATGCTAAAGCCTACGAAACCGAGAAGCACAATGTGTGG
 GCCACACATGAGTGCCTCCCTACTGACCCAAACCCCGAGGAAATCCACCTGGAGAATGTGAC
 AGAGGAATTCAACATGTGGAAAAACAATATGGTGGAGCAGATGCATACTGACATCATTGAGC
 TGTGGGATCAGTCTCTGAAGCCTTGCGTGAAACTGACACCACTGTGCGTCACTCTGCAGTGT
 ACTAACGTGACCAACAATATCACCGACGATATGCGGGCGAACTGAAGAATTGTTCTTTCAA
 CATGACCACAGAGCTGCGGGACAAGAAACAGAAAGTGACAGTCTGTTTTATAGACTGGATG

TGGTCCAGATCAATGAAAACCAGGGGAATCGAAGTAACAATTCAAACAAGGAGTACCGGCTG
ATCAATTGCAACACAAGTGCAATTACTCAGGCCTGTCCAAAAGTGTCAATTTGAACCTATCCC
AATTCATTATTGCGCTCCCGCAGGCTTCGCCATCCTGAAGTGTAAGATAAGAAGTTCAACG
GCACCGGGCCATGCCCTTCAGTGAGCACCGTCCAGTGACACACGGAATTAAGCCAGTGGTC
TCCACACAGCTGCTGCTGAATGGCTCTCTGGCTGAGGAAGAAGTGATGATCAGGAGCGAGAA
CATTACCAACAATGCAAAGAATATCCTGGTGCAGTTCAACACACCCGTCAGATTAATTGCA
CCCGCCCTAACAATAACACACGGAAATCTATCAGAATTGGACCCGGCCAGGCATTTTATGCC
ACCGGCGACATCATTGGGGATATCAGACAGGCACACTGTAATGTGAGCAAGGCCACTTGAA
CGAGACCCTGGGGAAGGTGGTCAAACAGCTGAGAAAGCATTTTCGAAACAACACTATCATCA
GGTTTGCCAATAGCTCCGGCGGGGACCTGGAAGTGACTACCCACAGCTTCAACTGCGGAGGC
GAGTTCTTTTACTGTAACACAAGTGGCCTGTTTAATTCAACCTGGATCAGCAACACATCCGT
GCAGGGATCCAATTCTACAGGCTCTAACGATAGTATCACTCTGCCCTGCCGCATTAAGCAGA
TCATTAATATGTGGCAGCGAATCGGGCAGGCTATGTATGCACCCCTATCCAGGGAGTGATT
AGATGTGTCAGCAATATCACTGGCCTGATTCTGACCCGCGACGGGGGATCAACTAACAGCAC
AACTGAGACCTTCCGGCCTGGAGGAGGAGACATGAGGGATAACTGGCGCTCCGAACTGTACA
AGTATAAAGTGGTCAAGATCGAGCCACTGGGAGTGGCACCAACCAGATGCAAACGAAGAGTG
GTCGGAAGGCGACGACGAAGAAGGGCAGTGGGAATTGGAGCTGTCTTCCTGGGCTTTCTGGG
GGCCGCTGGATCTACAATGGGGGAGCCAGTATGACCCTGACAGTCCAGGCTCGAAATCTGC
TGTCAGGAATCGTGCAGCAGCAGAGCAACCTGCTGAGAGCCCCTGAAGCTCAGCAGCATCTG
CTGAAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCTAGGGTGTGGCAGTGCAGCGCTA
CCTGCGAGATCAGCAGCTGCTGGGCATCTGGGGGTGTTCCGGAAAGCTGATTTGCTGTACCA
ATGTGCCATGGAACCTAGTTGGAGCAATCGCAACCTGTCCGAAATCTGGGACAATATGACT
TGGCTGCAGTGGGATAAGGAGATTAGCAACTACCCAGATCATCTACGGCCTGCTGGAAGA
GTCCCAGAATCAGCAGGAGAAGAACGAGCAGGACCTGCTGGCCCTGGAT

[00164] BG505_SOSIP_D664_MD39C_mC highest melting temp 82.5C

[00165] AENLWVTVYYGVVPVWKDAETTLFCASDAKAYETEKHNVWATHECVPTDPNPQEIH
LENVTEEFNMWKNMVEQMHTDI IELWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELK
NCSFNMTELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVS
FEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEEEVI
IRSENITNNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNVS
KATWNETLGKVVKQLRKHFGNNTIIRFAQSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWI
SNTSVQGSNSTGSNDSITLPCRICKIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGG
STNSTTETFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVS
LGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKLTWVGIKQLQARVL
AVEHYLRDQQLLGIWCGSKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
GLLEESQNQQEKNEQDLLALD

[00166] GCCGAGAATCTGTGGGTGACCGTGACTATGGCGTGCCCGTGTGGAAGGACGCAG
AGACCACACTGTTCTGCGCCAGCGATGCCAAGGCCTACGAGACAGAGAAGCACAACGTGTGG

GCCACCCACGAGTGCCTACAGACCCAAACCCCCAGGAGATCCACCTGGAGAATGTGAC
CGAGGAGTTTAACATGTGGAAGAACAATATGGTGGAGCAGATGCACACAGACATCATCGAGC
TGTGGGATCAGTCCCTGAAGCCCTGCGTGAAGCTGACCCCTCTGTGCGTGACACTGCAGTGT
ACCAACGTGACAAACAATATCACCGACGATATGAGGGGCGAGCTGAAGAATTGTAGCTTCAA
CATGACCACAGAGCTGCGGGACAAGAAGCAGAAGGTGTACTCCCTGTTTTATAGACTGGATG
TGGTGCAGATCAATGAGAACCAGGGCAATAGGTCTAACAATAGCAACAAGGAGTACCGCCTG
ATCAATTGCAACACCTCTGCCATCACACAGGCCTGTCCTAAGGTGAGCTTCGAGCCTATCCC
AATCCACTATTGCGCCCCAGCCGGCTTCGCCATCCTGAAGTGTAAGGATAAGAAGTTTAACG
GCACCGGCCCATGCCCTTCCGTGTCTACCGTGCAGTGTACACACGGCATCAAGCCTGTGGTG
AGCACACAGCTGCTGCTGAATGGCTCCCTGGCCGAGGAGGAAGTGATCATCCGGAGCGAGAA
CATCACCAACAATGCCAAGAATATCCTGGTGCAGCTGAACACACCAGTGCAGATCAATTGCA
CCAGGCCCCAACAATAACACAGTGAAGTCTATCCGCATCGGCCAGGCCAGGCCTTTTACTAT
ACCGGCGACATCATCGGCGACATCAGACAGGCCCACTGTAATGTGAGCAAGGCCACCTGGAA
CGAGACACTGGGCAAGGTGGTGAAGCAGCTGCGGAAGCACTTCGGCAATAACACCATCATCA
GATTTGCACAGAGCAGCGGAGGCGACCTGGAGGTGACCACACACTCCTTCAATTGCGGCGGC
GAGTTCTTTTACTGTAACACATCCGGCCTGTTTAATTCTACCTGGATCTCCAACACATCTGT
GCAGGGCAGCAATTCCACCGGCAGCAACGATTCCATCACACTGCCATGCAGGATCAAGCAGA
TCATCAACATGTGGCAGAGGATCGGACAGGCAATGTATGCACCACCTATCCAGGGCGTGATC
AGATGCGTGAGCAATATCACCGGCCTGATCCTGACACGCGACGGAGGCTCTACCAACAGCAC
CACAGAGACATTAGGCCCCGGCGGAGGCGACATGAGGGATAACTGGAGATCTGAGCTGTACA
AGTATAAGGTGGTGAAGATCGAGCCTCTGGGAGTGGCACCAACCAGGTGCAAGAGGAGAGTG
GTGGGCAGGCGCCGGAGAAGGCGCGCAGTGGGCATCGGAGCCGTGAGCCTGGGCTTTCTGGG
AGCAGCAGGCAGCACAAATGGGCGCAGCCAGCATGACCCTGACAGTGCAGGCCCGGAATCTGC
TGTCCGGCATCGTGCAGCAGCAGTCTAACCTGCTGAGAGCCCCAGAGCCCCAGCAGCACCTG
CTGAAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCAGAGTGCTGGCCGTGGAGCACTA
CCTGAGAGATCAGCAGCTGCTGGGCATCTGGGGATGCTCCGGCAAGCTGATCTGCTGTACCA
ATGTGCCCTGGAACCTAGCTGGAGCAATAGAAACCTGTCCGAGATCTGGGACAATATGACC
TGGCTGCAGTGGGATAAGGAGATCTCCAACCTACACACAGATCATCTATGGCCTGCTGGAGGA
GTCTCAGAATCAGCAGGAGAAGAACGAGCAGGACCTGCTGGCCCTGGAT

[00167] BG505_SOSIP_D664_MD9_mC improved yield/stability

[00168] AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQEIH
LENVTEEFNMWKNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELK
NCSFNMTELRLDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVS
FEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEFEV
IRSENITNNAKNILVQFNTPVQINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNVS
KATWNETLGKVVKQLRKHFGNNTIIRFANSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWI
SNTSVQGSNSTGSNDSITLPCRICKIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGG
STNSTTETFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRAVGIGAVF

LGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPERQQHLLKDTVWGIKQLQARVL
AVEHYLRDQQLLGIWGC SGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
GLLEESQNQQEKNEQDLLALD

[00169] GCTGAAAACCTGTGGGTCACTGTCTACTACGGCGTGCCTGTCTGGAAGGATGCTG
AACTACTCTGTTCTGTGCCTCTGATGCTAAAGCCTACGAAACCGAGAAGCACAATGTGTGG
GCCACACATGCTTGCGTCCCTACTGACCCAAACCCCCAGGAAATCCACCTGGAGAATGTGAC
AGAGGAATTCAACATGTGGAAAAACAATATGGTGGAGCAGATGCATACTGACATCATTTCCC
TGTGGGATCAGTCTCTGAAGCCTTGCGTGAACTGACACCACTGTGCGTCACTCTGCAGTGT
ACTAACGTGACCAACAATATCACCGACGATATGCGGGGCGAACTGAAGAATTGTTCTTTCAA
CATGACCACAGAGCTGCGGGACAAGAAACAGAAAGTGACAGTCTGTTTTATAGACTGGATG
TGGTCCAGATCAATGAAAACAGGGGAATCGAAGTAACAATTCAAACAAGGAGTACCGGCTG
ATCAATTGCAACACAAGTGCAATTACTCAGGCCTGTCCAAAAGTGTCAATTTGAACCTATCCC
AATTCATTATTGCGCTCCCGCAGGCTTCGCCATCCTGAAGTGTAAGATAAGAAGTTCAACG
GCACCGGGCCATGCCCTTCAGTGAGCACCGTCCAGTGACACACGGAATTAAGCCAGTGGTC
TCCACACAGCTGCTGCTGAATGGCTCTCTGGCTGAGGAAGAAGTGATGATCAGGAGCGAGAA
CATTACCAACAATGCAAAGAATATCCTGGTGCAGTTCAACACACCCGTCCAGATTAATTGCA
CCCGCCCTAACAATAACACACGGAAATCTATCAGAATTGGACCCGGCCAGGCATTTTATGCC
ACCGGCGACATCATTGGGGATATCAGACAGGCACACTGTAATGTGAGCAAGGCCACTTGAA
CGAGACCTGGGGAAGGTGGTCAAACAGCTGAGAAAGCATTTCGGAAACAACACTATCATCA
GGTTTGCCAATAGCTCCGGCGGGGACCTGGAAGTGACTACCCACAGCTTCAACTGCGGAGGC
GAGTTCTTTTACTGTAACACAAGTGGCCTGTTTAATTCAACCTGGATCAGCAACACATCCGT
GCAGGGATCCAATTCTACAGGCTCTAACGATAGTATCACTCTGCCCTGCCGCATTAAGCAGA
TCATTAATATGTGGCAGCGAATCGGGCAGGCTATGTATGCACCCCTATCCAGGGAGTGATT
AGATGTGTCAGCAATATCACTGGCCTGATTCTGACCCGCGACGGGGGATCAACTAACAGCAC
AACTGAGACCTTCCGGCCTGGAGGAGGAGACATGAGGGATAACTGGCGCTCCGAACTGTACA
AGTATAAAGTGGTCAAGATCGAGCCACTGGGAGTGGCACCAACCAGATGCAAACGAAGAGTG
GTCGGAAGGCGACGACGAAGAAGGGCAGTGGGAATTGGAGCTGTCTTCCTGGGCTTTCTGGG
GGCCGCTGGATCTACAATGGGGGCAGCCAGTATGACCCTGACAGTCCAGGCTCGAAATCTGC
TGTCAGGAATCGTGACGAGCAGAGCAACCTGCTGAGAGCCCCTGAACGGCAGCAGCATCTG
CTGAAGGACACCGTGTGGGGCATCAAGCAGCTGCAGGCTAGGGTGCTGGCAGTGCAGCACTA
CCTGCGAGATCAGCAGCTGCTGGGCATCTGGGGGTGTTCCGGAAAGCTGATTTGCTGTACCA
ATGTGCCATGGAACCTAGTTGGAGCAATCGCAACCTGTCCGAAATCTGGGACAATATGACT
TGGCTGCAGTGGGATAAGGAGATTAGCAACTACACCCAGATCATCTACGGCCTGCTGGAAGA
GTCCCAGAATCAGCAGGAGAAGAACGAGCAGGACCTGCTGGCCCTGGAT

[00170] BG505_SOSIP_D664_MD2_mC improved yield

[00171] AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIH
LENVTEEFNMWKNNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELK
NCSFNMTTELKDKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVS

FEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAESEVM
IRSENITNNAKNILVQFNTPVQINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNVS
KATWNETLGKVVKQLRKHFNGNTIIRFANSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWI
SNTSVQGSNSTGSNDSITLPCRICKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGG
STNSTTETFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVF
LGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQHLLKDTVWGIKQLQARVL
AVERYLRDQQLLGIWGC SGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
GLLEESQNNQEKNEQDLLALD

[00172] GCTGAAAACCTGTGGGTCACTGTCTACTACGGCGTGCCTGTCTGGAAGGATGCTG
AACTACTCTGTTCTGTGCCTCTGATGCTAAAGCCTACGAAACCGAGAAGCACAATGTGTGG
GCCACACATGCTTGCGTCCCTACTGACCCAAACCCCCAGGAAATCCACCTGGAGAATGTGAC
AGAGGAATTCAACATGTGGAAAAACAATATGGTGGAGCAGATGCATACTGACATCATTTCCC
TGTGGGATCAGTCTCTGAAGCCTTGCGTGAACTGACACCACTGTGCGTCACTCTGCAGTGT
ACTAACGTGACCAACAATATCACCGACGATATGCGGGGCGAACTGAAGAATTGTTCTTTCAA
CATGACCACAGAGCTGCGGGACAAGAAACAGAAAGTGACAGTCTGTTTTATAGACTGGATG
TGGTCCAGATCAATGAAAACAGGGGAATCGAAGTAACAATTCAAACAAGGAGTACCGGCTG
ATCAATTGCAACACAAGTGCAATTACTCAGGCCTGTCCAAAAGTGTCATTTGAACCTATCCC
AATTCAATTATTGCGCTCCCGCAGGCTTCGCCATCCTGAAGTGTAAGATAAGAAGTTCAACG
GCACCGGGCCATGCCCTTCAGTGAGCACCGTCCAGTGACACACGGAATTAAGCCAGTGGTC
TCCACACAGCTGCTGCTGAATGGCTCTCTGGCTGAGGAAGAAGTGATGATCAGGAGCGAGAA
CATTACCAACAATGCAAAGAATATCCTGGTGCAGTTCAACACACCCGTCCAGATTAATTGCA
CCCGCCCTAACAATAACACACGGAATCTATCAGAATTGGACCCGGCCAGGCATTTTATGCC
ACCGGCGACATCATTGGGGATATCAGACAGGCACACTGTAATGTGAGCAAGGCCACTTGAA
CGAGACCTTGGGGAAGGTGGTCAAACAGCTGAGAAAGCATTTCGGAAACAACACTATCATCA
GGTTTGCCAATAGCTCCGGCGGGGACCTGGAAGTGACTACCCACAGCTTCAACTGCGGAGGC
GAGTTCTTTTACTGTAACACAAGTGGCCTGTTTAATTCAACCTGGATCAGCAACACATCCGT
GCAGGGATCCAATTCTACAGGCTCTAACGATAGTATCACTCTGCCCTGCCGCATTAAGCAGA
TCATTAATATGTGGCAGCGAATCGGGCAGGCTATGTATGCACCCCTATCCAGGGAGTGATT
AGATGTGTCAGCAATATCACTGGCCTGATTCTGACCCGCGACGGGGGATCAACTAACAGCAC
AACTGAGACCTTCCGGCCTGGAGGAGGAGACATGAGGGATAACTGGCGCTCCGAACTGTACA
AGTATAAAGTGGTCAAGATCGAGCCACTGGGAGTGGCACCAACCAGATGCAAACGAAGAGTG
GTCGGAAGGCGACGACGAAGAAGGGCAGTGGGAATTGGAGCTGTCTTCCTGGGCTTTCTGGG
GGCCGCTGGATCTACAATGGGGGCAGCCAGTATGACCCTGACAGTCCAGGCTCGAAATCTGC
TGTCAGGAATCGTGACGAGCAGAGCAACCTGCTGAGAGCCCCTGAAGCTCAGCAGCATCTG
CTGAAGGACACCGTGTGGGGCATCAAGCAGCTGCAGGCTAGGGTGCTGGCAGTCCAGCGCTA
CCTGCGAGATCAGCAGCTGCTGGGCATCTGGGGGTGTTCCGGAAGCTGATTTGCTGTACCA
ATGTGCCATGGAACCTAGTTGGAGCAATCGCAACCTGTCCGAAATCTGGGACAATATGACT

TGGCTGCAGTGGGATAAGGAGATTAGCAACTACACCCAGATCATCTACGGCCTGCTGGAAGA
GTCCCAGAATCAGCAGGAGAAGAACGAGCAGGACCTGCTGGCCCTGGAT

[00173] # Computational designed hydrophobic core under the variable loops V1,V2,V3.
(apolarV1V2 + MD39)

[00174] BG505_SOSIP_D664_olio6_mC

[00175] AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQEIH
LENVTEEFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNVNTNNITDDMRGELK
NCSFNMTTELKQKQKVYSLFWRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVS
FEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEEVI
IRSENITNNAKNILVQLNTPVQINCTAPNNFTVKSIRIGPGQAFYYMGDIIGDIRQAHCNVS
KATWNETLGKVVQQLRKHFGNNTIIRFAQSSGGDLEVTHSFNCGGMFFYCNTSGLFNSTWI
SNTSVQGSNSTGSNDSITLPCRILIIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGG
STNSTTETFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVS
LGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARVL
AVEHYLRDQQLLGIWCGSKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
GLLEESQNQQEKNEQDLLALD

[00176] GCCGAAAACCTGTGGGTGACTGTCTACTATGGCGTGCCCGTCTGGAAGGACGCAG
AGACCACACTGTTCTGCGCTAGCGATGCTAAGGCCTACGAGACCGAGAAACACAACGTGTGG
GCAACCCATGCCTGCGTGCCTACAGACCCAAATCCCCAGGAAATCCACCTGGAGAACGTGAC
CGAGGAGTTCAACATGTGGAAGAACAATATGGTGGAACAGATGCATGAGGACATCATTTCCC
TGTGGGATCAGTCTCTGAAGCCCTGCGTGAACTGACCCCTCTGTGCGTCACACTGCAGTGT
ACAAACGTGACTAACAATATCACCGACGATATGCGGGGCGAACTGAAGAACTGTTCTTTCAA
TATGACTACCGAGCTGCGCGACAAGAAACAGAAAGTGACAGTCTGTTTTGGCGACTGGATG
TGGTCCAGATCAACGAAAATCAGGGGAACCGGAGTAACAACCTCAAATAAGGAGTATAGACTG
ATCAACTGCAATACCACTGCCATTACACAGGCTTGTCCTAAAGTGTCATTTCGAGCCTATCCC
AATTCATTACTGCGCCCCAGCTGGCTTCGCCATCCTGAAGTGTAAGATAAGAAGTTCAACG
GCACCGGGCCATGCCCTTCAGTGAGCACTGTCCAGTGTACCCACGGAATTAAGCCTGTGGTC
TCCACACAGCTGCTGCTGAACGGCTCTCTGGCTGAGGAAGAAGTGATCATTAGATCCGAGAA
TATCACTAACAATGCAAAGAACATTCTGGTGACAGCTGAATACCCAGTCCAGATCAACTGCA
CTGCCCCCAACAATTTACCGTGAAATCTATCCGGATTGGACCAGGCCAGGCTTTTTACTAT
ATGGGCGACATCATTGGGGATATTAGACAGGCACACTGTAACGTGAGCAAGGCCACATGGAA
TGAAACTCTGGGGAAGGTGGTCAAACAGCTGCGAAAAACATTTGGAACAATACAATCATTC
GATTTGCACAGAGCAGCGGAGGGGACCTGGAGGTGACAACTCACAGCTTCAACTGCGGAGGC
ATGTTCTTTTATTGTAATACTAGTGGCCTGTTTAACTCAACTTGGATCAGCAATACCTCCGT
GCAGGGCAGCAACAGCACCGGCTCTAATGATAGTATCACACTGCCATGCAGAATTAAGCTGA
TCATTAATATGTGGCAGAGGATCGGGCAGGCTATGTACGCACCCCCTATCCAGGGAGTGATT
CGGTGCGTGAGCAACATCACAGGCCTGATTCTGACTAGAGACGGGGGATCAACAAATAGCAC
CACAGAGACTTTTCAGGCCCGGCGGAGGAGACATGCGAGATAACTGGCGATCCGAACTGTACA

AGTATAAAGTGGTCAAGATCGAGCCTCTGGGAGTGGCACCAACCCGATGCAAACGAAGAGTG
 GTCGGGAGGCGCCGACGGAGAAGGGCTGTGGGGATTGGAGCAGTCTCTCTGGGCTTTCTGGG
 GGCCGCTGGATCTACAATGGGGGACAGCCAGTATGACTCTGACCGTCCAGGCAAGGAACCTGC
 TGTCAAGGAATCGTGCAGCAGCAGAGCAATCTGCTGCGCGCTCCAGAACCCCAGCAGCACCTG
 CTGAAGGACACCCATTGGGGCATCAAGCAGCTGCAGGCAAGGGTGCTGGCAGTTCGAGCACTA
 CCTGCGAGATCAGCAGCTGCTGGGCATCTGGGGATGCAGCGGAAAGCTGATTTGCTGTACAA
 ACGTGCCCTGGAACAGCAGCTGGAGCAACCGGAATCTGTCCGAAATCTGGGACAACATGACA
 TGGCTGCAGTGGGATAAGGAGATTAGCAATTACACTCAGATCATCTACGGCCTGCTGGAAGA
 GTCCCAGAACCAGCAGGAGAAGAACGAGCAGGACCTGCTGGCCCTGGAT

[00177] # Core mutations from mammalian display data to improve apex binding built onto olio6

[00178] BG505_SOSIP_D664_MD53_mC

[00179] AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQEIH
 LENVT EEFNMWKNMVEQM HEDI I SLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELK
 NCSFNMTTEL RDKKQKVYSLFWRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVS
 FEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVLSTQLLLNGLSLAE EEEVI
 VRSENITNNAKNILVQLNTPVQINCTAPNFTVKSIRIGPGQAFYYMGDIIGDIRQAHCNVS
 KATWNETLGKVVQQLRKHFGNNTIIRFAQSSGGDLEVTTTHSFNCGGMFFFCNTSGLFNSTWI
 SNTSVQGSNSTGSNDSITLPCRILKLIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGG
 STNSTTETFRPGGDMRDVWRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVS
 LGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARVL
 AVEHYLRDQQLLGIWCGSKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
 GLLEESQNQQEKNEQDLLALD

[00180] GCCGAGAACCTGTGGGTGACCGTGTACTATGGCGTGCCCGTGTGGAAGGACGCAG
 AGACCACACTGTTTTGCGCCAGCGATGCCAAGGCCTATGAGACAGAGAAGCACAACGTGTGG
 GCAACCCACGCATGCGTGCCTACAGACCCAAATCCCCAGGAGATCCACCTGGAGAACGTGAC
 CGAGGAGTTCAATATGTGGAAGAACAATATGGTGGAGCAGATGCACGAGGACATCATCTCCC
 TGTGGGATCAGTCTCTGAAGCCCTGCGTGAAGCTGACCCCTCTGTGCGTGACACTGCAGTGT
 ACCAACGTGACAAACAATATCACCGACGATATGCGGGGCGAGCTGAAGAACTGTTCTTTTAA
 TATGACCACAGAGCTGAGGGACAAGAAGCAGAAGGTGTACAGCCTGTTCTGGCGCCTGGATG
 TGGTGCAGATCAACGAGAATCAGGGCAACAGGTCTAACAATAGCAATAAGGAGTATCGCCTG
 ATCAACTGCAATACCTCCGCCATCACACAGGCCTGTCCTAAGGTGTCTTTTGAGCCTATCCC
 AATCCACTACTGCGCCCCAGCCGGCTTTGCCATCCTGAAGTGTAAGGATAAGAAGTTCAACG
 GCACCGGCCCATGCCCTTCCGTGTCTACCGTGCAGTGTACACACGGCATCAAGCCTGTGCTG
 TCTACACAGCTGCTGCTGAACGGCAGCCTGGCCGAGGAGGAAGTGATCGTGCGGAGCGAGAA
 TATCACCAACAATGCCAAGAACATCCTGGTGCAGCTGAATACACCAGTGCAGATCAACTGCA
 CCGCCCCCAACAACCTTACCGTGAAGTCCATCCGGATCGGCCAGGCCAGGCCTTCTACTAT
 ATGGGCGACATCATCGGCGACATCAGACAGGCCCACTGTAACGTGTCTAAGGCCACCTGGAA

TGAGACACTGGGCAAGGTGGTGAAGCAGCTGAGGAAGCACTTTGGCAACAATACCATCATCA
 GGTTTCGCACAGAGCAGCGGAGGCGACCTGGAGGTGACCACACACTCCTTTAACTGCGGCGGC
 ATGTTCTTTTCTGTAAATACAAGCGGCCTGTTCAACTCCACCTGGATCTCCAATACATCTGT
 GCAGGGCAGCAACTCCACCGGCAGCAATGATTCCATCACACTGCCATGCCGGATCAAGCTGA
 TCATCAATATGTGGCAGAGAATCGGCCAGGCCATGTATGCACCACCTATCCAGGGCGTGATC
 AGATGCGTGAGCAACATCACCGGCCTGATCCTGACAAGAGACGGCGGCTCTACCAATAGCAC
 CACAGAGACCTTCCGGCCCCGGCGGAGGCGACATGAGGGACGTGTGGAGATCCGAGCTGTACA
 AGTATAAGGTGGTGAAGATCGAGCCTCTGGGAGTGGCACCAACCAGGTGCAAGAGGAGAGTG
 GTGGGCAGGCGCCGGAGAAGGCGCGCAGTGGGCATCGGCGCCGTGTCTCTGGGCTTCTGGG
 AGCAGCAGGCAGCACAAATGGGCGCAGCCTCTATGACCCTGACAGTGCAGGCCAGGAACCTGC
 TGAGCGGCATCGTGCAGCAGCAGTCCAATCTGCTGCGCGCCCCAGAGCCACAGCAGCACCTG
 CTGAAGGACACCCACTGGGGCATCAAGCAGCTGCAGGCCAGAGTGCTGGCCGTGGAGCACTA
 CCTGAGAGATCAGCAGCTGCTGGGCATCTGGGGATGCTCCGGCAAGCTGATCTGCTGTACCA
 ACGTGCCCTGGAACAGCAGCTGGTCTAACCGGAATCTGAGCGAGATCTGGGACAACATGACC
 TGGCTGCAGTGGGATAAGGAGATCAGCAATTACACACAGATCATCTATGGCCTGCTGGAGGA
 GTCCCAGAACCAGCAGGAGAAGAATGAGCAGGACCTGCTGGCCCTGGAT

[00181] BG505_SOSIP_D664_MD37

[00182] AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQEIH
 LENVTEEFNMWKNMVEQM**H**EDIISLWDQSLKPC**C**KLTPLCVTLQCTNVTNNITDDMRGELK
 NCSFNMTTELKDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTS**A**CTQACPKVS
 FEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAE**E**VI
 IRSENITNNAKNILVQLNTPVQINCTRPNNNTRKSIRIGPG**C**AFYATGDIIGDIRQAHCNVS
 KATWNETLGKVVQQLRKHFNGNTIIRFA**Q**SSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWI
 SNTSVQGSNSTGSNDSITLPCRIKQIINMWQRIGQ**C**MYAPPIQGVIRCVSNITGLILTRDGG
 STNSTTETFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRAVGIGAV**S**
 LGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPE**PQ****EHLHKDTH**WGIKQLQARVL
 AVEHYLRDQQLLGIWCGSKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
 GLEESQ**Q**QEKNEQDLLALD

[00183] D

[00184] (mutations relative to BG505 SOSIP in bold, underline)

[00185] # MD37 has two extra disulfide bonds to stabilize the V3 loop (V120C, Q315C) and to prevent CD4 induced conformational changes (I201C, A433C).

[00186] IV: Examples of trimers with combined germline-targeting mutations and stabilization mutations

[00187] #This construct includes germline-targeting mutations built on top of the MD39 stabilization mutations from III

[00188] BG505_SOSIP_D664_MD39_9mut2A_mC (MD39+9MUT)

[00189] AENLWVTVYYGVVPVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQEIH
LENVTEEFNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNY**APFLIN**NMRGELK
NCSFNMTELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVS
FEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGLAE~~EEVI~~
IRSENITNNAKNILVQLNTPVQINCTRPNNNT**VKS**IRIGPGQAFY**Y**TGDIIGDIR**MAH**CNVS
KATWNETLGKVVKQLRKHFGNNTIIRFA**Q**SSGGDLEVTTHSFNCGGEFFYCNTSGLFNSTWI
SNTSVQGSNSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGG
STNSTTETFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAV**S**
LGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPE**PQ**QHLL**KD**THWGIKQLQARVL
AVEHYLRDQQLLGIWGC SGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
GLLEESQNQQEKNEQDLLALD

[00190] GCCGAGAATCTGTGGGTGACCGTGACTATGGCGTGCCCGTGTGGAAGGACGCAG
AGACCACACTGTTCTGCGCCAGCGATGCCAAGGCCTACGAGACAGAGAAGCACAACGTGTGG
GCAACCCACGCATGCGTGCTACAGACCCAAACCCCGAGGAGATCCACCTGGAGAATGTGAC
CGAGGAGTTTAACATGTGGAAGAACAATATGGTGAGCAGATGCACGAGGACATCATCTCCC
TGTGGGATCAGTCTCTGAAGCCCTGCGTGAAAGCTGACCCCTCTGTGCGTGACCCTGCAGTGT
ACAACTATGCCCCCTTCTGATCAACAATATGAGGGCGAGCTGAAGAATTGTTCTTTTAA
CATGACCACAGAGCTGCGGGACAAGAAGCAGAAAGTGACAGCCTGTTCTATAGACTGGATG
TGGTGAGATCAATGAGAACCAGGGCAATAGGTCTAACAATAGCAACAAGGAGTACCGCCTG
ATCAATTGCAACACCTCCGCCATCACACAGGCCTGTCTAAGGTGTCTTTTGAGCCTATCCC
AATCCACTATTGCGCCCCAGCCGGCTTCGCCATCCTGAAGTGTAAGGATAAGAAGTTTAACG
GCACCGGCCCATGCCCTTCCGTGTCTACCGTGAGTGACACACGGCATCAAGCCTGTGGTG
TCTACACAGCTGCTGCTGAATGGCAGCCTGGCCGAGGAGGAAGTGATCATCCGGAGCGAGAA
CATCACCAACAATGCCAAGAATATCCTGGTGAGCTGAACACACAGTGAGATCAATTGCA
CCAGGCCCAACAATAACACAGTGAAGTCCATCCGCATCGGCCAGGCCAGGCCTTCTACTAT
ACCGGCGACATCATCGGCGACATCAGAATGGCCCACTGTAACGTGAGCAAGGCCACCTGGAA
CGAGACACTGGGCAAGGTGGTGAAGCAGCTGCGGAAGCACTTCGGCAATAACACCATCATCA
GATTTGCACAGAGCAGCGGAGGCGACCTGGAGGTGACCACACACTCCTTCAATTGCGGCGGC
GAGTTCTTTTACTGTAACACAAGCGGCCTGTTTAATTCCACCTGGATCTCCAACACATCTGT
GCAGGGCAGCAATTCCACCGGCAGCAACGATTCCATCACACTGCCATGCAGGATCAAGCAGA
TCATCAACATGTGGCAGAGGATCGGACAGGCAATGTATGCACCACCTATCCAGGGCGTGATC
AGATGCGTGAGCAATATCACCGGCCTGATCCTGACACGCGACGGAGGCTCTACCAACAGCAC
CACAGAGACATTAGGCCCGGCGGAGGCGACATGAGGGATAACTGGAGATCCGAGCTGTACA
AGTATAAGGTGGTGAAGATCGAGCCTCTGGGAGTGGCACCAACCAGGTGCAAGAGGAGAGTG
GTGGGCAGGCGCCGGAGAAGGCGCGCAGTGGGCATCGGAGCCGTGTCTCTGGGCTTTCTGGG
AGCAGCAGGCTCCACAATGGGCGCAGCCTCTATGACCCTGACAGTGACAGGCCCGGAATCTGC
TGAGCGGCATCGTGACAGCAGTCCAACCTGCTGAGAGCCCCAGAGCCCCAGCAGCACCTG
CTGAAGGACACCCACTGGGGCATCAAGCAGCTGCAGGCCAGAGTGCTGGCCGTGGAGCACTA

CCTGAGAGATCAGCAGCTGCTGGGCATCTGGGGATGCTCCGGCAAGCTGATCTGCTGTACCA
 ATGTGCCTTGGAACCTAGCTGGTCTAATAGAAACCTGAGCGAGATCTGGGACAATATGACC
 TGGCTGCAGTGGGATAAGGAGATCAGCAACTACACACAGATCATCTATGGCCTGCTGGAGGA
 GTCCCAGAATCAGCAGGAGAAGAACGAGCAGGACCTGCTGGCCCTGGAT

[00191] #This construct includes germline-targeting mutations built on top of the MD39
 stabilization mutations from III

[00192] BG505_SOSIP_D664_MD39_10mut_mC

[00193] AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIH
 LENVTEEFNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNY**APFL**INDMRGELK
 NCSFNMTELDRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVS
 FEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAE~~EEVI~~
 IRSENITNNAKNILVQLNTPVQINCTRPNNNT**VKS**IRIGPGQAFY**YFGDI**IGDIR**MAH**CNVS
 KATWNETLGKVVKQLRKHFGNNTIIRFA**QSS**GGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWI
 SNTSVQGSNSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGG
 STNSTTETFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRAVGIGAV**S**
 LGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPE**PQQHLLKDTH**WGIKQLQARVL
 AVE**HY**LRDQQLLGIWCGSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
 GLLEESQNQQEKNEQDLLALD

[00194] GCCGAGAATCTGTGGGTGACCGTGTACTATGGCGTGCCCGTGTGGAAGGACGCAG
 AGACCACACTGTTCTGCGCCAGCGATGCCAAGGCCTACGAGACAGAGAAGCACAACGTGTGG
 GCAACCCACGCATGCGTGCCTACAGACCCAAACCCCCAGGAGATCCACCTGGAGAATGTGAC
 CGAGGAGTTTAACATGTGGAAGAACAATATGGTGGAGCAGATGCACGAGGACATCATCTCCC
 TGTGGGATCAGTCTCTGAAGCCCTGCGTGAAGCTGACCCCTCTGTGCGTGACCCTGCAGTGT
 ACAAACTATGCCCCCTTCCTGATCAACGATATGAGGGGCGAGCTGAAGAATTGTTCTTTTAA
 CATGACCACAGAGCTGCGGGACAAGAAGCAGAAAGTGACAGCCTGTTCTATAGACTGGATG
 TGGTGCAGATCAATGAGAACCAGGGCAATAGGTCTAACAATAGCAACAAGGAGTACCGCCTG
 ATCAATTGCAACACCTCCGCCATCACACAGGCCTGTCCCTAAGGTGTCTTTTGAGCCTATCCC
 AATCCACTATTGCGCCCCAGCCGGCTTCGCCATCCTGAAGTGTAAGGATAAGAAGTTTAAACG
 GCACCGGCCCATGCCCTTCGTGTCTACCGTGCAGTGTACACACGGCATCAAGCCTGTGGTG
 TCTACACAGCTGCTGCTGAATGGCAGCCTGGCCGAGGAGGAAGTGATCATCCGGAGCGAGAA
 CATCACCAACAATGCCAAGAATATCCTGGTGCAGCTGAACACACCAGTGCAGATCAATTGCA
 CCAGGCCCAACAATAACACAGTGAAGTCCATCCGCATCGGCCAGGCCAGGCCTTCTACTAT
 TTTGGCGACATCATCGGCGACATCAGAATGGCCCACTGTAACGTGAGCAAGGCCACCTGGAA
 CGAGACACTGGGCAAGGTGGTGAAGCAGCTGCGGAAGCACTTCGGCAATAACACCATCATCA
 GATTTGCACAGAGCAGCGGAGGCGACCTGGAGGTGACCACACACTCCTTCAATTGCGGCGGC
 GAGTTCTTTTACTGTAACACAAGCGGCCTGTTTAATTCCACCTGGATCTCCAACACATCTGT
 GCAGGGCAGCAATTCCACCGGCAGCAACGATTCCATCACACTGCCATGCAGGATCAAGCAGA
 TCATCAACATGTGGCAGAGGATCGGACAGGCAATGTATGCACCACCTATCCAGGGCGTGATC

AGATGCGTGAGCAATATCACCGGCCTGATCCTGACACGCGACGGAGGCTCTACCAACAGCAC
CACAGAGACATTTCAGGCCCCGGCGGAGGCGACATGAGGGATAACTGGAGATCCGAGCTGTACA
AGTATAAGGTGGTGAAGATCGAGCCTCTGGGAGTGGCACCAACCAGGTGCAAGAGGAGAGTG
GTGGGCAGGCGCCGGAGAAGGCGCGCAGTGGGCATCGGAGCCGTGTCTCTGGGCTTTCTGGG
AGCAGCAGGCTCCACAATGGGCGCAGCCTCTATGACCCTGACAGTGCAGGCCCGGAATCTGC
TGAGCGGCATCGTGCAGCAGCAGTCCAACCTGCTGAGAGCCCCAGAGCCCCAGCAGCACCTG
CTGAAGGACACCCACTGGGGCATCAAGCAGCTGCAGGCCAGAGTGCTGGCCGTGGAGCACTA
CCTGAGAGATCAGCAGCTGCTGGGCATCTGGGGATGCTCCGGCAAGCTGATCTGCTGTACCA
ATGTGCCTTGGAACCTCTAGCTGGTCTAATAGAAACCTGAGCGAGATCTGGGACAATATGACC
TGGCTGCAGTGGGATAAGGAGATCAGCAACTACACACAGATCATCTATGGCCTGCTGGAGGA
GTCCCAGAATCAGCAGGAGAAGAACGAGCAGGACCTGCTGGCCCTGGAT

[00195] #This construct includes germline-targeting mutations built on top of the MD39 stabilization mutations from III

[00196] BG505_SOSIP_D664_MD39_10mut2A_mC

[00197] AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIH
LENVTEEFNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNY**APFLIN**NMRGELK
NCSFNMTTEL RDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVS
FEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAE~~EEVI~~
IRSENITNNAKNILVQLNTPVQINCTRPNNNT**VKS**SIRIGPGQAFYY**FGDI**IGDIR**MAH**CNVS
KATWNETLGKVVKQLRKHFGNNTIIRFA**QSS**GGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWI
SNTSVQGSNSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGG
STNSTTETFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAV**S**
LGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPE**PQQHLLKDTHW**GIKQLQARVL
AVE**HYLRDQQL**LGIWGCSGLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
GLLEESQNQQEKNEQDLLALD

[00198] GCCGAGAATCTGTGGGTGACCGTGACTATGGCGTGCCCGTGTGGAAGGACGCAG
AGACCACACTGTTCTGCGCCAGCGATGCCAAGGCCTACGAGACAGAGAAGCACAACGTGTGG
GCAACCCACGCATGCGTGCCTACAGACCCAAACCCCCAGGAGATCCACCTGGAGAATGTGAC
CGAGGAGTTTAACATGTGGAAGAACAATATGGTGGAGCAGATGCACGAGGACATCATCTCCC
TGTGGGATCAGTCTCTGAAGCCCTGCGTGAAGCTGACCCCTCTGTGCGTGACCCTGCAGTGT
ACAAACTATGCCCCCTTCCTGATCAACAATATGAGGGGCGAGCTGAAGAATTGTTCTTTTAA
CATGACCACAGAGCTGCGGGACAAGAAGCAGAAAGTGACAGCCTGTTCTATAGACTGGATG
TGGTGCAGATCAATGAGAACCAGGGCAATAGGTCTAAACAATAGCAACAAGGAGTACCGCCTG
ATCAATTGCAACACCTCCGCCATCACACAGGCCTGTCCTAAGGTGTCTTTTGAGCCTATCCC
AATCCACTATTGCGCCCCAGCCGGCTTCGCCATCCTGAAGTGTAAGGATAAGAAGTTTAACG
GCACCGGCCCATGCCCTTCGGTGTCTACCGTGCAGTGTACACACGGCATCAAGCCTGTGGTG
TCTACACAGCTGCTGCTGAATGGCAGCCTGGCCGAGGAGGAAGTGATCATCCGGAGCGAGAA
CATCACCAACAATGCCAAGAATATCCTGGTGCAGCTGAACACACCAGTGCAGATCAATTGCA

CCAGGCCCAACAATAACACAGTGAAGTCCATCCGCATCGGCCCAGGCCAGGCCTTCTACTAT
 TTTGGCGACATCATCGGCGACATCAGAATGGCCCACTGTAACGTGAGCAAGGCCACCTGGAA
 CGAGACACTGGGCAAGGTGGTGAAGCAGCTGCGGAAGCACTTCGGCAATAACACCATCATCA
 GATTTGCACAGAGCAGCGGAGGCGACCTGGAGGTGACCACACACTCCTTCAATTGCGGCGGC
 GAGTTCTTTTACTGTAACACAAGCGGCCTGTTTAATTCCACCTGGATCTCCAACACATCTGT
 GCAGGGCAGCAATTCCACCGGCAGCAACGATTCCATCACACTGCCATGCAGGATCAAGCAGA
 TCATCAACATGTGGCAGAGGATCGGACAGGCAATGTATGCACCACCTATCCAGGGCGTGATC
 AGATGCGTGAGCAATATCACCGGCCTGATCCTGACACGCGACGGAGGCTCTACCAACAGCAC
 CACAGAGACATTTCAGGCCCCGGCGGAGGCGACATGAGGGATAACTGGAGATCCGAGCTGTACA
 AGTATAAGGTGGTGAAGATCGAGCCTCTGGGAGTGGCACCAACCAGGTGCAAGAGGAGAGTG
 GTGGGCAGGCGCCGGAGAAGGCGCGCAGTGGGCATCGGAGCCGTGTCTCTGGGCTTTCTGGG
 AGCAGCAGGCTCCACAATGGGCGCAGCCTCTATGACCCTGACAGTGCAGGCCCCGGAATCTGC
 TGAGCGGCATCGTGCAGCAGCAGTCCAACCTGCTGAGAGCCCCAGAGCCCCAGCAGCACCTG
 CTGAAGGACACCCACTGGGGCATCAAGCAGCTGCAGGCCAGAGTGCTGGCCGTGGAGCACTA
 CCTGAGAGATCAGCAGCTGCTGGGCATCTGGGGATGCTCCGGCAAGCTGATCTGCTGTACCA
 ATGTGCCTTGGAACCTAGCTGGTCTAATAGAAACCTGAGCGAGATCTGGGACAATATGACC
 TGGCTGCAGTGGGATAAGGAGATCAGCAACTACACACAGATCATCTATGGCCTGCTGGAGGA
 GTCCCAGAATCAGCAGGAGAAGAACGAGCAGGACCTGCTGGCCCTGGAT

[00199] #This construct includes germline-targeting mutations built on top of the MD39 stabilization mutations from III

[00200] BG505_SOSIP_D664_MD39_11mut2A_mC

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLNVTEE
 FNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNY**APNLLS**NMRGELKNCSFNMT
 TELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIH
 YCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEFEV**I**IRSENIT
 NNAKNILVQLNTPVQINCTRPNNNT**V**KSIRIGPGQAF**YFYF**GDIIGDIR**MA**HCVNSKATWNET
 LGKVVKQLRKHFNGNT**I**IRFA**Q**SSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQG
 SNSTGSNDS**I**VLPCRIKQ**I**INMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTE
 TFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRAVGIGAV**SL**GLFLGAA
 GSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPE**PQQHLLKD**THWGIKQLQARVLAVE**HY**LR
 DQQLLGIWGC**SG**KLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQ**I**IYGLLEESQ
 NQ**Q**EKNEQDLLALD

[00201] GCCGAAAACCTGTGGGTGACTGTCTATTATGGCGTGCCCGTGTGGAAAGATGCTG
 AA**A**CTACTCTGTTCTGTGCAAGCGATGCTAAGGCCTACGAGACAGAGAAGCACAACGTGTGG
 GCAACCCACGCATGCGTGCCTACAGACCCAAATCCCCAGGAGATCCACCTGGAGAACGTGAC
 AGAGGAGTTCAATATGTGGAAGAACAATATGGTGGAGCAGATGCACGAGGACATCATCTCCC
 TGTGGGATCAGTCTCTGAAGCCCTGCGTGAAGCTGACCCCTCTGTGCGTGACCCTGCAGTGT
 ACAAATTATGCCCCAAACCTGCTGTCCAATATGAGGGGCGAGCTGAAGAACTGTTCTTTCAA

TATGACCACAGAGCTGCGGGACAAGAAGCAGAAGGTGTACAGCCTGTTTTATAGACTGGATG
 TGGTGCAGATCAATGAGAACCAGGGCAACAGGTCTAAACAATAGCAATAAGGAGTACCGCCTG
 ATCAATTGCAACACCAGCGCCATCACACAGGCCTGTCCTAAGGTGTCCTTTGAGCCTATCCC
 AATCCACTATTGCGCCCCAGCCGGCTTCGCCATCCTGAAGTGTAAGGATAAGAAGTTTAAACG
 GCACCGGCCCATGCCCTTCCGTGTCTACAGTGCAGTGTACACACGGCATCAAGCCAGTGGTG
 TCTACACAGCTGCTGCTGAACGGCAGCCTGGCCGAGGAGGAAGTGATCATCCGGAGCGAGAA
 TATCACCACAATGCCAAGAACATCCTGGTGCAGCTGAATACACCAGTGCAGATCAACTGCA
 CCAGGCCCAACAATAACACAGTGAAGTCTATCCGCATCGGCCCCGGCCAGGCCTTCTACTAT
 TTTGGCGACATCATCGGCGATATCAGAATGGCCCACTGTAACGTGAGCAAGGCCACCTGGAA
 TGAGACACTGGGCAAGGTGGTGAAGCAGCTGCGGAAGCACTTCGGCAATAACACCATCATCA
 GATTTGCACAGTCCTCCGGCGGCGACCTGGAGGTGACCACACACTCCTTCAACTGCGGCGGC
 GAGTTCTTTTACTGTAATACCAGCGGCCTGTTTAACTCCACCTGGATCTCCAATACATCTGT
 GCAGGGCAGCAACTCCACAGGCAGCAATGATTCCATCGTGCTGCCCTGCAGGATCAAGCAGA
 TCATCAACATGTGGCAGCGCATCGGCCAGGCCATGTATGCCCCCTCCCATCCAGGGCGTGATC
 AGATGCGTGAGCAACATTACCGGCCTGATCCTGACAAGAGATGGCGGATCTACCAATAGCAC
 AACCGAGACATTCAGGCCCGGCGGCGGCGACATGAGAGATAACTGGAGATCTGAGCTGTACA
 AGTATAAGGTGGTGAAGATTGAGCCTCTGGGAGTGGCACCACAAGATGCAAGAGAAGAGTG
 GTGGGCAGGCGCCGGAGAAGGCGCGCAGTGGGCATTGGAGCCGTGTCCCTGGGCTTTCTGGG
 AGCAGCAGGATCCACAATGGGAGCAGCCTCTATGACCCTGACAGTGCAGGCCCGGAACCTGC
 TGAGCGGCATCGTGACAGCAGTCCAATCTGCTGAGAGCCCCAGAGCCCCAGCAGCACCTG
 CTGAAGGACACCCACTGGGGCATCAAGCAGCTGCAGGCACGGGTGCTGGCAGTGGAGCACTA
 CCTGAGAGATCAGCAGCTGCTGGGCATCTGGGGCTGTAGCGGCAAGCTGATCTGCTGTACCA
 ACGTGCCCTGGAATTCTAGCTGGTCTAATAGAAACCTGAGCGAGATCTGGGACAACATGACC
 TGGCTGCAGTGGGATAAGGAGATCTCCAATTACACACAGATCATCTATGGCCTGCTGGAGGA
 ATCACAGAATCAGCAGGAAAAGAACGAACAGGATCTGCTGGCACTGGAT

[00202] V: Trimers with modified surfaces or of different strains than BG505, that can be employed in strategic boosting regimens

[00203] # Variable loop cocktails. Shown on MD39 background

[00204] BG505_SOSIP_MD39_VLC1-03

[00205] AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQEIH
 LENVTEEFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTDWDNATLANMTGEIK
 NCSFNMTELRDKKQKVYSLFYELDIPIENYISNNNTSNTSYRLINCNTSAITQACPKVS
 FEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEFEVI
 IRSENITNNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIIGDIRQAHCNVS
ATQWEQTLKGIAAKLLEHFGNNTIIRFAQSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNGSSW
NLNKTKENTTNLENGTITLPCRIKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGG
NKSAGIETFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVS
 LGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARVL

AVEHYLRDQQLLGIWCGSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
GLLEESQNQQEKNEQDLLALD

[00206] GCCGAGAACCTGTGGGTGACCGTGACTATGGCGTGCCCGTGTGGAAGGACGCAG
AGACCACACTGTTCTGCGCCTCCGATGCCAAGGCCTACGAGACAGAGAAGCACAACGTGTGG
GCAACCCACGCATGCGTGCCTACAGACCCAAATCCCCAGGAGATCCACCTGGAGAATGTGAC
AGAGGAGTTTAACATGTGGAAGAACAATATGGTGGAGCAGATGCACGAGGACATCATCTCCC
TGTGGGATCAGTCTCTGAAGCCCTGCGTGAAGCTGACCCCTCTGTGCGTGACCCTGCAGTGT
ACAGACTGGGATAATGCCACCCTGGCCAACATGACAGGCGAGATCAAGAATTGTTCTTCAA
CATGACCACAGAGCTGCGGGACAAGAAGCAGAAGGTGTACTCTCTGTTTTATGAGCTGGACA
TCATCCCCATCGAGAACGAGTACATCAGCAACAATAACACCTCTAATACAAGCTATAGACTG
ATCAACTGCAATACCTCTGCCATCACACAGGCCTGTCCTAAGGTGAGCTTCGAGCCTATCCC
AATCCACTATTGCGCCCCAGCCGGCTTCGCCATCCTGAAGTGTAAGGATAAGAAGTTTAACG
GCACCGGCCCATGCCCTAGCGTGTCCACCGTGCGTGTACACACGGCATCAAGCCTGTGGTG
AGCACACAGCTGCTGCTGAACGGCTCCCTGGCCGAGGAGGAAGTGATCATCAGGAGCGAGAA
TATCACCAATAACGCCAAGAATATCCTGGTGCAGCTGAACACACCAGTGCAGATCAATTGCA
CCCGGCCCAATAACAATACAGTGAAGTCCATCAGAATCGGCCCAGGCCAGGCCTTTTACTAT
ACCGGCGACATCATCGGCGACATCAGGCAGGCCCACTGTAACGTGTCTGCCACCCAGTGGGA
GCAGACACTGAAGGGCATCGCCGCCAAGCTGCTGGAGCACTTCGGCAACAATACCATCATCA
GGTTTGCACAGAGCAGCGGAGGCGACCTGGAGGTGACCACACACTCCTTCAATTGCGGCGGC
GAGTTCTTTTACTGTAACACCTCTGGCCTGTTTAATGGCTCTAGCTGGAACCTGAATAAGAC
AAAGGAGAACACCACAAATCTGGAGAACGGCACCATCACACTGCCATGCAGGATCAAGCAGA
TCATCAACATGTGGCAGAGGATCGGACAGGCAATGTATGCACCACCTATCCAGGGCGTGATC
AGATGCGTGAGCAACATCACCGGCCTGATCCTGACAAGAGATGGCGGCAATAAGAGCGCCGG
CATCGAGACCTTCCGGCCCCGGCGGAGGCGACATGAGGGATAACTGGAGATCCGAGCTGTACA
AGTATAAGGTGGTGAAGATCGAGCCTCTGGGCGTGGCCCCAACAAAGATGCAAGAGGAGAGTG
GTGGGCAGGCGCCGGAGAAGGCGCGCAGTGGGCATCGGAGCCGTGAGCCTGGGCTTTCTGGG
AGCAGCAGGCAGCACAATGGGCGCAGCCAGCATGACCCTGACAGTGCAGGCCAGGAACCTGC
TGTCCGGCATCGTGACGAGCAGTCTAATCTGCTGCGCGCCCCAGAGCCACAGCAGCACCTG
CTGAAGGACACCCACTGGGGCATCAAGCAGCTGCAGGCCAGGGTGCTGGCAGTGGAGCACTA
CCTGAGGGATCAGCAGCTGCTGGGCATCTGGGGATGCAGCGGCAAGCTGATCTGCTGTACCA
ATGTGCCTTGGAACCTCTCTTGAGCAACCGGAATCTGTCCGAGATCTGGGACAACATGACC
TGGCTGCAGTGGGATAAGGAGATCAGCAATTACACACAGATCATCTATGGCCTGCTGGAGGA
GTCCCAGAATCAGCAGGAGAAGAACGAGCAGGACCTGCTGGCCCTGGAT

[00207] BG505_SOSIP_MD39_VLC2-04

[00208] AENLWTVYYGVFVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQEIH
LENVTEEFNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCS**SDYEGN**TTRQ**NI**TMKE
EKGEIKNCSFNMTTELRLDKKQKVYSLFY**KLDITPI**EED**NN**S**NN**SS**SAN**SS**SN**AN**Y**T**NY**RLI
NCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVS

TQLLNGLSLAEEVVIIRSENITNNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPGQAFYYT
GDIIGDIRQAHCNVSGTKWKNTLKQIVKKLGDHFGNNTIIRFAQSSGGDLEVTTHSFNCGGE
FFYCNTSGLFNSTWNRNNGTWNDVEGLNYTNGNDTITLPCRIKQIINMWQRIGQAMYAPPIQ
GVIRCVSNITGLILTRDGGNDTDKNETFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCK
RRVVGRRRRRRRAVGIGAVSLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQ
QHLLKDTHWGIKQLQARVLAVEHYLRDQQLLGIWGC SGKLI CCTNVPWNSSWSNRNLSEIWD
NMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALD

[00209] GCCGAGAACCTGTGGGTGACCGTGTACTATGGCGTGCCCGTGTGGAAGGACGCAG
AGACCACACTGTTCTGCGCCTCTGATGCCAAGGCCTACGAGACAGAGAAGCACAACTGTGG
GCAACCCACGCATGCGTGCCTACAGACCCAAACCCCCAGGAGATCCACCTGGAGAACGTGAC
CGAGGAGTTTAATATGTGGAAGAACAATATGGTGGAGCAGATGCACGAGGACATCATCTCTC
TGTGGGATCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCTCTGTGCGTGACACTGCAGTGT
AGCGACTATGAGGGCAACACCACACGGCAGAATATCACCATGAAGGAGGAGAAGGGCGAGAT
CAAGAACTGTTCTTTCAATATGACCACAGAGCTGAGAGATAAGAAGCAGAAGGTGTACAGCC
TGTTTTTATAAGCTGGACATCACCCCATCGAGGAGGATAACAATTCCAACAATAGCTCCTCT
GCCAATAGCTCCAATTCTAACGCCAATTACACAACTATAGGCTGATCAACTGCAATACCTC
TGCCATCACACAGGCCTGTCTAAGGTGAGCTTCGAGCCTATCCCAATCCACTACTGCGCCC
CAGCCGGCTTCGCCATCCTGAAGTGTAAGGATAAGAAGTTTAACGGCACCGGCCCATGCCCT
AGCGTGTCCACCGTGCAGTGTACACACGGCATCAAGCCTGTGGTGAGCACACAGCTGCTGCT
GAATGGCTCCCTGGCCGAGGAGGAAGTGATCATCCGCTCCGAGAACATACCAACAATGCCA
AGAACATCCTGGTGCAGCTGAATACACCAGTGCAGATCAACTGCACCCGGCCCAACAATAAC
ACAGTGAAGTCCATCAGAATCGGCCCAGGCCAGGCCTTTTACTATACCGGCGACATCATCGG
CGACATCCGGCAGGCCCACTGTAACTGAGCGGCACCAAGTGGAAGAACACACTGAAGCAGA
TCGTGAAGAAGCTGGGCGACCACTTCGGCAATAACACCATCATCAGATTTGCCAGTCTAGC
GGCGGCGATCTGGAGGTGACCACACACAGCTTCAATTGCGGCGGCGAGTTCTTTTACTGTAA
TACCTCCGGCCTGTTTAACTCTACATGGAACCGGACCAATGGCACATGGAATGACGTGGAGG
GCCTGAACTATAACCAACGGCAATGATACCATCACACTGCCATGCAGGATCAAGCAGATCATC
AATATGTGGCAGAGGATCGGACAGGCAATGTACGCACCACCTATCCAGGGCGTGATCAGATG
CGTGAGCAACATCACCGGCCTGATCCTGACAAGAGACGGCGGCAACGACACCGGATAAGAATG
AGACATTCAGGCCCCGGGAGGCGACATGAGGGATAACTGGAGATCCGAGCTGTACAAGTAT
AAGGTGGTGAAGATCGAGCCTCTGGGAGTGGCACCAACCAGGTGCAAGAGGAGAGTGGTGGG
CAGGCGCCGGAGAAGGCGCGCAGTGGGCATCGGAGCCGTGTCCCTGGGCTTTCTGGGAGCAG
CAGGCTCTACAATGGGCGCAGCCAGCATGACCCTGACAGTGCAGGCCAGGAATCTGCTGTCC
GGCATCGTGCAGCAGCAGTCTAACCTGCTGCGCGCCCCAGAGCCACAGCAGCACCTGCTGAA
GGACACCCACTGGGGCATCAAGCAGCTGCAGGCCAGGGTGCTGGCAGTGGAGCACTATCTGC
GCGATCAGCAGCTGCTGGGCATCTGGGGATGCAGCGGCAAGCTGATCTGCTGTACAAACGTG
CCCTGGAACAGCAGCTGGAGCAACAGGAATCTGTCCGAGATCTGGGACAATATGACCTGGCT

GCAGTGGGATAAGGAGATCAGCAACTACACACAGATCATCTATGGCCTGCTGGAGGAGTCCC
AGAACCAGCAGGAGAAGAATGAGCAGGACCTGCTGGCCCTGGAT

[00210] BG505_SOSIP_MD39_VLC2-08

[00211] AENLWVTVYYGVVPWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIH
LENVTEEFNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCVTLKNCNSNCSISRN
ISIEMDGEIKNCSFNMTELRLDKKQKVYSLFYRLDIVPIESSNNSQLSNNSQVSNNSSQSSNY
SQYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHG
IKPVVSTQLLNGSLAEEEVIIIRSENITNNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPG
QAFYYTGDIIIGDIRQAHCNVSKKDWEKTLQQVATKLGQHFGNNTIIRFAQSSGGDLEVTTHS
FNCGGEFFYCNTSGLFNSTWIRNSSNSTWNSSASNSTELNSNITLPCRIKQIINMWQRIGQA
MYAPPIQGVIRCVSNITGLILTRDGG**HETENKT**TETFRPGGGDMRDNRSELYKYKVVKIEPL
GVAPTRCKRRRVGRRRRRRRAVGIGAVSLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSN
LLRAPEPQQHLLKDTHWGIKQLQARVLAVEHYLRDQQLGIWGCSGKLI CCTNVPWNSSWSN
RNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALD

[00212] GCCGAGAACCTGTGGGTGACAGTGTACTATGGCGTGCCCGTGTGGAAGGACGCAG
AGACCACACTGTTCTGCGCCAGCGATGCCAAGGCCTACGAGACCGAGAAGCACAACGTGTGG
GCAACACACGCATGCGTGCCTACCGACCCAAATCCCCAGGAGATCCACCTGGAGAATGTGAC
CGAGGAGTTTAACATGTGGAAGAACAATATGGTGGAGCAGATGCACGAGGACATCATCTCCC
TGTGGGATCAGTCTCTGAAGCCCTGCGTGAAGCTGACCCCTCTGTGCGTGACACTGCAGTGC
GTGACCCCTGAAGAACTGCAGCAATTCCAAGTGTCTATCAGCAGGAATATCTCCATCGAGAT
GGATGGCGAGATCAAGAATTGTTCTTTCAACATGACCACAGAGCTGCGGGACAAGAAGCAGA
AAGTGACAGCCTGTTCTACCGGCTGGACATCGTGCCCATCGAGAGCAGCAACAATTCTCAG
CTGAGCAACAATTCCCAGGTGTCTAACAATAGCCAGTCTAGCAACTACTCCCAGTATCGCCT
GATCAATTGCAACACCTCTGCCATCACACAGGCCTGTCTTAAGGTGAGCTTCGAGCCTATCC
CAATCCACTATTGCGCCCCAGCCGGCTTCGCCATCCTGAAGTGTAAGGACAAGAAGTTTAAC
GGCACAGGCCCCCTGCCCTTCCGTGTCTACCGTGCAGTGTACACACGGCATCAAGCCTGTGGT
GTCTACCCAGCTGCTGCTGAACGGCAGCCTGGCAGAGGAGGAAGTGATCATCCGGAGCGAGA
ATATCACAAACAATGCCAAGAATATCCTGGTGCAGCTGAACACCCCAGTGCAGATCAATTGC
ACACGGCCCCAACAATAACACCGTGAAGTCCATCAGAAATCGGCCAGGCCAGGCCTTTTACTA
TACAGGCGACATCATCGGCGACATCCGGCAGGCCCACTGTAACGTGTCTAAGAAGGACTGGG
AGAAGACACTGCAGCAGGTGGCCACCAAGCTGGGCCAGCACTTCGGCAATAACACCATCATC
AGATTTGCCAGTCTCTGGCGGCGATCTGGAGGTGACCACACACTCCTTCAATTGCGGCGG
CGAGTTCTTTTACTGTAACACAAGCGGCCTGTTAATTCACCTGGATCAGAAATAGCTCCA
ACTCTACCTGGAACAGCAGCGCCAGCAACTCCACAGAGCTGAATAGCAACATCACCCCTGCCA
TGCAGGATCAAGCAGATCATCAACATGTGGCAGAGGATCGGACAGGCAATGTATGCACCACC
TATCCAGGGCGTGATCAGATGCGTGAGCAATATCACAGGCCTGATCCTGACCCGCGATGGAG
GACACGAGACCGAGAACAAGACCGAGACATTCAGGCCCGGCGGAGGCGACATGAGGGATAAT
TGGAGATCCGAGCTGTACAAGTATAAGGTGGTGAAGATCGAGCCTCTGGGAGTGGCACCAAC

AAGGTGCAAGAGGAGAGTGGTGGGCAGGCGCCGAGAAGGCGCGCAGTGGGCATCGGAGCCG
 TGAGCCTGGGCTTTCTGGGAGCAGCAGGCAGCACAAATGGGCGCAGCCTCTATGACCCTGACA
 GTGCAGGCCCGGAACCTGCTGAGCGGCATCGTGCAGCAGCAGTCCAATCTGCTGAGAGCCCC
 AGAGCCCCAGCAGCACCTGCTGAAGGACACACACTGGGGCATCAAGCAGCTGCAGGCCAGGG
 TGCTGGCAGTGGAGCACTACCTGAGGGATCAGCAGCTGCTGGGCATCTGGGGATGCTCCGGC
 AAGCTGATCTGCTGTACCAATGTGCCTTGGAACCTCTTGGTCTAATCGGAACCTGAGCGA
 GATCTGGGACAACATGACATGGCTGCAGTGGGATAAGGAGATCAGCAATTACACCCAGATCA
 TCTATGGCCTGCTGGAGGAGTCCCAGAATCAGCAGGAGAAGAACGAGCAGGACCTGCTGGCC
 CTGGAT

[00213] BG505_SOSIP_MD39_VLC3-13

[00214] AENLWTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIH
 LENVT EEFNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCT**TNAALTNVTITNGPNI**
TEEIRNCSFNMTTELRDKKQKVYSLFY**KLDLVQINGSGGEYRL**INCNTSAITQACPKVSFEP
 IPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAE EEEVIIRS
 ENITNNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNV**SGTK**
 WNETL**KQVAGKLRKH**FGNNTIIRFAQSSGGDLEVTTHSFNCGGEFFYCNTSGLF**NSTWPENG**
TMEGSNGTITLPCRICKIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGG**NSTTD**TET
 FRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRRVGRRRRRRRAVGIGAVSLGFLGAAG
 STMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARVLAVEHYLRD
 QQLLGIWGCSGKLI CCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQN
 QQEKNEQDLLALD

[00215] GCCGAGAACCTGTGGGTGACCGTGACTATGGCGTGCCAGTGTGGAAGGACGCCG
 AGACCACACTGTTCTGCGCCTCTGATGCCAAGGCCTACGAGACAGAGAAGCACAACTGTGG
 GCAACCCACGCATGCGTGCCAACAGACCCAAACCCCGAGGAGATCCACCTGGAGAATGTGAC
 CGAGGAGTTTAACATGTGGAAGAACAATATGGTGGAGCAGATGCACGAGGACATCATCAGCC
 TGTGGGATCAGTCCCTGAAGCCCTGCGTGAAGCTGACCCCTCTGTGCGTGACCCCTGCAGTGT
 ACAAATGCCGCCCTGACCAACGTGACCATCACAAATGGCCCCAACATCACAGAGGAGATCAG
 GAATTGTTCTTCAACATGACCACAGAGCTGCGCGACAAGAAGCAGAAGGTGTACTCTCTGT
 TTTATAAGCTGGATCTGGTGCAGATCAATGGCAGCGGCGGCGAGTACCGGCTGATCAATTGC
 AACACCAGCGCCATCACACAGGCCTGTCCTAAGGTGTCCTTCGAGCCTATCCCAATCCACTA
 TTGCGCCCCAGCCGGCTTCGCCATCCTGAAGTGTAAGGACAAGAAGTTTAACGGCACCCGGCC
 CATGCCCTTCCGTGTCTACCGTGAGTGTACACACGGCATCAAGCCTGTGGTGTCCACACAG
 CTGCTGCTGAATGGCTCTCTGGCCGAGGAGGAAGTGATCATCAGGAGCGAGAACATCACCAA
 CAATGCCAAGAATATCCTGGTGCAGCTGAACACACCAAGTGCAGATCAATTGCACCCGGCCCA
 ACAATAACACAGTGAAGTCTATCAGAATCGGCCCAGGCCAGGCCTTTTACTATACCGGCGAC
 ATCATCGGCGACATCCGCCAGGCCCACTGTAATGTGAGCGGCACCAAGTGGAACGAGACACT
 GAAGCAGGTGGCCGGCAAGCTGAGGAAGCACTTCGGCAATAACACCATCATCCGCTTTGCAC
 AGAGCAGCGGAGGCGATCTGGAGGTGACCACACACTCCTTCAATTGCGGCGGCGAGTTCTTT

TACTGTAACACATCTGGCCTGTTTAATAGCACCTGGCCCGAGAACGGCACAATGGAGGGCTC
TAATGGCACCATCACACTGCCTTGCCGGATCAAGCAGATCATCAACATGTGGCAGAGAATCG
GCCAGGCCATGTATGCACCACCTATCCAGGGCGTGATCAGATGCGTGAGCAATATCACCGGC
CTGATCCTGACAAGAGACGGCGGCAACTCCACCACAGATACCGAGACATTCCGGCCCCGGCGG
AGGCGACATGAGGGATAACTGGAGAAGCGAGCTGTACAAGTATAAGGTGGTGAAGATCGAGC
CTCTGGGCGTGGCCCCAACCAGATGCAAGAGGAGAGTGGTGGGCAGGCGCCGGAGAAGGCGC
GCAGTGGGCATCGGAGCCGTGTCCCTGGGCTTTCTGGGAGCAGCAGGCAGCACAATGGGCGC
AGCCTCCATGACCCTGACAGTGCAGGCCAGGAATCTGCTGTCTGGCATCGTGCAGCAGCAGA
GCAACCTGCTGCGCGCCCCAGAGCCACAGCAGCACCTGCTGAAGGACACCCACTGGGGCATC
AAGCAGCTGCAGGCCAGGGTGCTGGCAGTGGAGCACTACCTGAGGGATCAGCAGCTGCTGGG
CATCTGGGGATGCTCCGGCAAGCTGATCTGCTGTACCAATGTGCCTTGGAACCTCTAGCTGGT
CCAATAGGAACCTGTCTGAGATCTGGGACAATATGACCTGGCTGCAGTGGGATAAGGAGATC
TCCAACTACACACAGATCATCTATGGCCTGCTGGAGGAGTCTCAGAATCAGCAGGAGAAGAA
CGAGCAGGACCTGCTGGCCCTGGAT

[00216] BG505_SOSIP_VLC1-03_DS21_mC

[00217] AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQEIH
LENVTEEFNMWKNMVEQMHTDIISLWDQSLKPCCKLTPLCVTLQCTDWDNATLANMTGEIK
NCSFNMTTEL RDKKQKVYSLFYELDIIP IENEYISNNNTSNTSYRLINCNTSAITQACPKVS
FEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAE E EVM
IRSENITNNAKNILVQFNTPVQINCTRPNNNTRKSIRIGPGCAFYATGDIIGDIRQAHCNVS
ATQWEQTLKGIAAKLLEHFGNNTIIRFANSSGGDLEVTHSFNCGGEFFYCNTSGLFNGSSW
NLNKT KENTTNLENGTITLPCR IKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGG
NKSAGIETFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVF
LGFLGAAGSTMGAASMTLT VQARNLLSGIVQQQSNLLRAPEAQQHLLKLT VWGIKQLQARVL
AVERYLRDQQLLGIWCGSKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
GLLEESQNQQEKNEQDLLALD

[00218] GCTGAAAATCTGTGGGTCACTGTGTATTATGGGGTGCCTGTGTGGAAAGACGCTG
AGACTACTCTGTTCTGTGCCTCCGACGCTAAAGCCTACGAAACCGAGAAGCACAACGTGTGG
GCAACCCATGCCTGCGTCCCTACAGACCCAAATCCCCAGGAAATCCACCTGGAGAATGTGAC
AGAGGAGTTCAACATGTGGAAAAACAATATGGTGGAGCAGATGCATACTGACATCATTAGCC
TGTGGGATCAGTCCCTGAAGCCTTGCTGCAAACCTGACCCCACTGTGCGTCACCCTGCAGTGT
ACAGACTGGGATAATGCTACCCTGGCAAACATGACAGGCGAAATCAAGAATTGTAGTTTCAA
CATGACCACAGAGCTGAGGGACAAGAAACAGAAAGTGTACTCCCTGTTTTATGAACTGGACA
TCATTCCCATCGAAAACGAGTACATCTCTAACAACAACACAAGTAACACTTCATATCGCCTG
ATCAACTGCAATACTAGCGCCATTACCCAGGCTTGTCCAAAGGTGTCCTTTGAGCCTATCCC
AATTCATTACTGCGCCCCCGCTGGCTTCGCCATCCTGAAGTGTAAGATAAGAAGTTCAACG
GCACCGGGCCATGCCCTTCCGTGTCTACAGTCCAGTGTACTCACGGAATTAAGCCAGTGGTC
AGTACACAGCTGCTGCTGAACGGCTCACTGGCAGAGGAAGAAGTGATGATCCGGAGCGAGAA

CATCACAAACAACGCTAAGAACATCCTGGTGCAGTTCAACACTCCCGTCCAGATTAATTGCA
 CTAGACCTAACAACAACACCCGGAAAAGCATCAGAATTGGACCCGGCTGCGCCTTTTATGCT
 ACCGGCGACATCATTGGCGACATCCGGCAGGCCCACTGTAACGTGAGCGCTACTCAGTGGGA
 ACAGACCCTGAAGGGGATTGCCGCTAACTGCTGGAGCATTTCGGAAACAATACCATCATT
 GATTTGCCAACAGCTCCGGCGGGGACCTGGAAGTGACTACCCACTCTTTCAATTGCGGAGGC
 GAGTTCTTTTACTGTAACACTAGTGGACTGTTTAATGGCTCTAGTTGGAACCTGAATAAGAC
 CAAAGAAAACACAATAATCTGGAGAACGGCACCATCACACTGCCCTGCCGAATTAAGCAGA
 TCATTAACATGTGGCAGCGGATCGGCCAGGCAATGTATGCCCCCCTATCCAGGGCGTGATC
 AGATGTGTCTCCAACATCACAGGACTGATTCTGACTAGGGATGGGGGAAACAAGTCTGCCGG
 GATCGAGACTTTTCAGGCCTGGCGGGGGAGACATGAGGGATAACTGGCGCTCCGAAGTGTACA
 AGTATAAAGTGGTCAAGATCGAGCCACTGGGAGTGGCTCCAACACGATGCAAACGAAGAGTG
 GTCGGAAGGCGACGACGAAGAAGGGCTGTGGGAATCGGAGCAGTCTTCCTGGGCTTTCTGGG
 GGCAGCCGGATCAACAATGGGCGCTGCAAGCATGACTCTGACCGTGCAGGCACGAAACCTGC
 TGTCCGGAATCGTCCAGCAGCAGTCTAATCTGCTGCGAGCTCCTGAAGCACAGCAGCATCTG
 CTGAAGCTGACAGTGTGGGGCATCAAGCAGCTGCAGGCACGGGTGCTGGCAGTCGAGCGCTA
 TCTGCGAGACCAGCAGCTGCTGGGCATCTGGGGGTGTTCTGGAAAGCTGATTTGCTGTACCA
 ATGTGCCCTGGAACAGCAGCTGGTCTAACAGGAATCTGAGTGAAATCTGGGACAACATGACC
 TGGCTGCAGTGGGATAAGGAGATTTCAAATTACACACAGATCATCTACGGCCTGCTGGAAGA
 GAGCCAGAATCAGCAGGAGAAGAACGAGCAGGACCTGCTGGCCCTGGAT

[00219] BG505_SOSIP_VLC2-04_DS21_mC

[00220] AENLWVTVYYGVVPVKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQEIH
 LENVT EEFNMWKNMVEQMHTDIISLWDQSLKPCCKLTPLCVTLQCSDYEGNTRQNTITMKE
 EKGEIKNCSFNMTTEL RDKKQKVYSLFYKLDITPI EEDNNSNNSSANSNSNANYTNYRLI
 NCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVS
 TQLLLNGSLAE E EVMIRSENITNNAKNILVQFNTVPVQINCTRPNNNTRKSIRIGPGCAFYAT
 GDIIGDIRQAHCNVSGTKWKNTLQKQIVKKLGDHFGNNTIIRFANSSGGDLEVTTHSFNCGGE
 FFYCNTSGLFNSTWNRTNGTWN DVEGLNYTNGNDTITLPCRIKQIINMWQRIGQAMYAPPIQ
 GVIRCVSNITGLILTRDGGNDTDKNETFRPGGDMRDNRSELYKYKVVKIEPLGVAPTRCK
 RRVVGRRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQ
 QHLLKLTWGIKQLQARVLAVERYLRDQQLGIWGC SGKLI CCTNVPWNSSWSNRNLSEIWD
 NMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALD

[00221] GCCGAAAATCTGTGGGTCACTGTGTATTATGGGGTCCCTGTGTGGAAGGATGCCG
 AAATACTCTGTTCTGTGCCTCCGACGCCAAAGCCTACGAAACCGAGAAGCACAATGTGTGG
 GCCACCCATGCTTGCGTCCCTACAGACCCAAACCCCAAGGAAATCCACCTGGAGAACGTGAC
 AGAGGAGTTCAACATGTGGAAAAACAATATGGTCGAACAGATGCATACTGACATCATTTCCC
 TGTGGGATCAGTCTCTGAAGCCTTGCTGCAAACCTGACACCACTGTGCGTCACTCTGCAGTGT
 TCCGACTATGAGGGCAACACACACGGCAGAATATCACCATGAAGGAGGAAAAAGGGGAAAT
 TAAGAACTGTTCAATATGACTACCGAGCTGAGAGATAAGAAACAGAAGGTGTACAGCC

TGTTTTATAAACTGGACATCACTCCCATTGAGGAAGATAACAATAGCAACAATAGCTCCTCT
GCAAATAGTTCAAATCCAACGCCAATTACACCAACTATAGGCTGATCAACTGCAATACCAG
TGCAATTACACAGGCCTGTCCAAAGGTGTCATTTGAGCCTATCCCAATTCATTACTGCGCTC
CCGCAGGCTTCGCCATCCTGAAGTGTAAGATAAGAAGTTCAACGGCACCGGGCCATGCCCT
TCAGTGAGCACTGTCCAGTGTACCCACGGGATCAAGCCAGTGGTCTCCACACAGCTGCTGCT
GAATGGATCTCTGGCTGAGGAAGAAGTGATGATCAGGAGCGAGAACATTACCAACAATGCAA
AGAACATCCTGGTGCAGTTCAATACACCCGTCCAGATTAAGTGCCTCGCCCTAACAAATAAC
ACCCGGAATCCATCAGAATTGGACCCGGCTGCGCATTTTATGCCACCGGCGACATCATTTGG
CGACATCAGGCAGGCCCACTGTAATGTGTCTGGCACCAAGTGGAAAAACACACTGAAGCAGA
TTGTGAAGAACTGGGAGACCATTTTCGGCAATAACACAATCATTCGCTTTTGCTAATAGCTCC
GGCGGGGATCTGGAAGTGACAACTCACAGCTTCAACTGCGGAGGCGAGTTCTTTTACTGTAA
TACAAGTGGCCTGTTTAACTCAACTTGAACCGCACAAATGGGACTTGAATGACGTGGAGG
GGCTGAACTATACCAACGGAAATGATACCATCACACTGCCCTGCCGAATTAAGCAGATCATT
AATATGTGGCAGCGGATCGGCCAGGCTATGTACGCACCCCCTATCCAGGGCGTGATCAGATG
TGTCAGTAACATCACTGGACTGATTCTGACCAGGGACGGGGGAAACGACACAGATAAAAATG
AGACTTTCCGGCCTGGCGGGGGAGACATGAGGGATAACTGGCGCTCTGAACTGTACAAGTAT
AAAGTGGTCAAGATCGAGCCACTGGGAGTGGCCCCACAAGATGCAAACGAAGAGTGGTCGG
AAGGCGACGACGAAGAAGGGCAGTGGGGATCGGAGCTGTCTTCTGGGCTTTCTGGGGGCCG
CTGGATCTACTATGGGAGCAGCCAGTATGACTCTGACCGTGCAGGCTCGCAATCTGCTGTCA
GGGATCGTCCAGCAGCAGAGCAACCTGCTGCGAGCCCCTGAAGCTCAGCAGCATCTGCTGAA
GCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCTCGGGTGCTGGCAGTCGAGCGCTACCTGC
GAGATCAGCAGCTGCTGGGCATCTGGGGGTGTAGCGGAAAGCTGATTTGCTGTACAAACGTG
CCCTGGAACAGCAGCTGGAGCAACAGAAATCTGTCCGAAATCTGGGACAATATGACTTGGCT
GCAGTGGGATAAGGAGATTTCTAACTACACCCAGATCATCTACGGCCTGCTGGAAGAGAGTC
AGAACCAGCAGGAGAAGAACGAGCAGGACCTGCTGGCCCTGGAT

[00222] BG505_SOSIP_VLC2-08_DS21_mC

[00223] AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIH
LENVTEEFNMWKNMVEQMHTDIISLWDQSLKPCCKLTPLCVTLQCVTLKNCSNSNCSISRN
ISIEMDGEIKNCSFNMTELRDKKQKVYSLFYRLDIVPIESSNNSQLSNNSQVSNNSSQSSNY
SQYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHG
IKPVVSTQLLNLGSLAEEVMIRSENITNNAKNILVQFNTVPVQINCTRPNNNTRKSIRIGPG
CAFYATGDIIGDIRQAHCNVSKKDWEKTLQQVATKLGQHFGNNTIIRFANSSGGDLEVTTHS
FNCGGEFFYCNTSGLFNSTWIRNSSNSTWNSSASNSTELNSNITLPCRIKQIINMWQRIGQA
MYAPPIQGVIRCVSNITGLILTRDGGHETENKTETFRPGGGDMRDNRSELYKYKVVKIEPL
GVAPTRCKRRVVGRRRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSN
LLRAPEAQQHLLKLTWVGIKQLQARVLAVERYLRDQQLGIWGC SGKLI CCTNPWNSSWSN
RNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALD

[00224] GCCGAAAATCTGTGGGTCACTGTGTATTATGGCGTGCCCGTGTGGAAGGATGCCG
AAACTACTCTGTTCTGTGCCTCCGATGCTAAAGCCTACGAAACCGAGAAGCACAACGTGTGG
GCCACTCATGCTTGCGTCCCTACCGACCCAAATCCCCAGGAAATCCACCTGGAGAATGTGAC
AGAGGAGTTCAACATGTGGAAAAACAATATGGTCGAGCAGATGCATACTGACATCATTAGCC
TGTGGGATCAGTCCCTGAAGCCTTGCTGCAAACCTGACCCCACTGTGCGTGACACTGCAGTGT
GTCACTCTGAAGAACTGCTCTAATAGTAACTGTTCAATCAGCAGGAATATCAGCATTGAAAT
GGATGGCGAGATTAAGAATTGTTCTTCAACATGACCACAGAACTGAGAGACAAGAAACAGA
AAGTGTACTCCCTGTTCTACCGGCTGGACATCGTCCCCATTGAGAGCTCCAACAATAGCCAG
CTGTCCAACAATTCTCAGGTGAGTAACAATTCACAGTCTAGTAACTACTCTCAGTATCGCCT
GATCAATTGCAACACTAGTGCAATTACCCAGGCCTGTCCAAAGGTGTCATTTGAGCCTATCC
CAATTCATTACTGCGCTCCCGCAGGCTTCGCCATCCTGAAGTGTAAGACAAGAAGTTCAAC
GGCACCGGGCCATGCCCTTCCGTGTCTACAGTCCAGTGTACTCACGGCATTAAAGCCAGTGGT
CAGCACACAGCTGCTGCTGAACGGGTCCCTGGCAGAGGAAGAAGTGATGATCAGGTCTGAGA
ATATTACCAACAATGCCAAGAATATCCTGGTGCAGTTCAACACACCCGTCCAGATTAATTGC
ACACGCCCTAACAATAACACTCGGAAATCTATCAGAATTGGACCCGGCTGCGCATTTTATGC
CACAGGCGACATCATTGGCGACATCAGGCAGGCCCACTGTAACGTGAGCAAGAAAGACTGGG
AGAAGACCCCTGCAGCAGGTGGCTACAAAACCTGGGACAGCATTTTCGGCAATAACACCATCATT
CGCTTTGCAAACCTCAAGCGGCGGGGATCTGGAAGTGACTACCCACAGCTTCAATTGCGGAGG
CGAGTTCTTTTACTGTAACACTTCTGGCCTGTTTAATAGTACCTGGATCAGAAACAGCAGCA
ACAGCACCTGGAATAGTTTCTAGTAGTAACTCAACAGAGCTGAACAGCAACATCACTCTGCCC
TGCCGAATTAAGCAGATCATTAAACATGTGGCAGCGGATCGGGCAGGCTATGTATGCACCCCC
TATCCAGGGAGTGATTGCTGTGTCTAGCAATATCACCGGCCTGATTCTGACACGAGACGGGG
GACATGAAACCGAGAACAAAACAGAGACTTTCCGGCCTGGCGGGGGAGACATGAGGGATAAT
TGGCGCTCCGAACCTGTACAAGTATAAAGTGGTCAAGATCGAGCCACTGGGGGTGGCCCCCAC
TAGATGCAAACGAGAGTGGTGCGAAGGCGACGACGAAGAAGGGCAGTGGGAATCGGAGCTG
TCTTCCTGGGCTTTCTGGGAGCAGCTGGCAGCACAATGGGCGCAGCCTCTATGACCCTGACA
GTGCAGGCTCGGAACCTGCTGAGTGGCATCGTCCAGCAGCAGTCAAATCTGCTGAGAGCCCC
TGAAGCTCAGCAGCACCTGCTGAAGCTGACAGTGTGGGGCATCAAGCAGCTGCAGGCTCGGG
TGCTGGCAGTCGAGCGCTATCTGCGAGATCAGCAGCTGCTGGGCATCTGGGGGTGTAGCGGA
AAGCTGATTTGCTGTACTAATGTGCCCTGGAACAGCAGCTGGTCAAATAGAAACCTGAGCGA
AATCTGGGACAACATGACTTGGCTGCAGTGGGATAAGGAGATTTCTAATTACACCCAGATCA
TCTACGGCCTGCTGGAAGAGAGTCAGAATCAGCAGGAGAAGAACGAGCAGGACCTGCTGGCC
CTGGAT

[00225] BG505_SOSIP_VLC3-13_DS21_mC

[00226] AENLWVTVYYGVVPVKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQEIH
LENVTEEFNMWKNMVEQMHTDIISLWDQSLKPCCKLTPLCVTLQCTNAALTNVTITNGPNI
TEEIRNCSFNMTTEL RDKKQKVYSLFYKLDLVQINGSGGEYRLINCNTSAITQACPKVSFEP
IPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEFEVMIRS

ENITNNAKNILVQFNTVPVQINCTRPNNNTRKSIRIGPGCAFYATGDIIGDIRQAHNCNVSGTK
 WNETLKQVAGKLRKHFGNNTIIRFANSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWPENG
 TMEGSNGTITLPCRKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGGNSTTDTET
 FRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRRVVGRRRRRRAVGIGAVFLGFLGAAG
 STMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQQHLLKLTWVGIKQLQARVLAVERYLRD
 QQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQN
 QQEKNEQDLLALD

[00227] GCCGAAAATCTGTGGGTGACTGTCTATTATGGGGTCCCTGTGTGGAAGGATGCCG
 AAATACTCTGTTCTGTGCCAGCGATGCTAAGGCCTACGAAACCGAGAAGCACAATGTGTGG
 GCAACCCATGCCTGCGTCCCTACAGACCCAAACCCCCAGGAAATCCACCTGGAGAATGTGAC
 CGAGGAGTTCAACATGTGGAAAAACAATATGGTGGAAACAGATGCATACAGACATCATTTCCC
 TGTGGGATCAGTCTCTGAAGCCTTGCTGCAAACCTGACCCCACTGTGCGTCACTCTGCAGTGT
 ACCAATGCCGCTCTGACCAACGTGACCATCACAAATGGCCCAAACATCACAGAGGAAATTCG
 CAATTGTTCTTTCAACATGACCACAGAGCTGCGAGACAAGAAACAGAAGGTGTACAGCCTGT
 TTTATAAACTGGATCTGGTCCAGATCAATGGGTCCGGCGGGGAATACAGGCTGATCAATTGC
 AACACAAGTGCTATTACTCAGGCATGTCCAAAGGTGTCATTTGAGCCTATCCCAATTCATTA
 TTGCGCCCCCGCTGGCTTCGCCATCCTGAAGTGTAAGACAAGAAGTTCAACGGAACCTGGCC
 CCTGCCCTTCAGTGAGCACCGTCCAGTGACACACGGCATTAAGCCAGTGGTCTCCACCCAG
 CTGCTGCTGAATGGGTCTCTGGCTGAGGAAGAAGTGATGATCCGGTCCGAGAACATTACTAA
 CAATGCAAAGAATATCCTGGTGCAGTTCAACACCCCCGTCCAGATTAATTGCACTAGACCTA
 ACAATAACACCCGGAATCCATCAGAATTGGGCCCGGATGCGCTTTTTATGCAACCGGGGAC
 ATCATTGGCGACATCCGCCAGGCCCACTGTAATGTGTCTGGCACTAAGTGGAACGAGACCCT
 GAAACAGGTGGCCGGCAAGCTGCGAAAAATTTTCGGGAATAACACAATCATTTCGGTTTGCTA
 ATAGCTCCGGAGGCGATCTGGAAGTGACTACCCACAGTTTCAACTGCGGGGGAGAGTTCTTT
 TACTGTAACTAGTGGACTGTTTAATTCAACATGGCCTGAAAACGGCACTATGGAGGGCAG
 CAATGGCACTATCACCCCTGCCATGCAGAATTAAGCAGATCATTAACATGTGGCAGAGGATCG
 GGCAGGCCATGTATGCTCCCCCTATCCAGGGAGTGATTTCGGTGTGTCTCAAATATCACAGGC
 CTGATTCTGACTAGAGACGGCGGGAACAGCACAACTGATACAGAGACTTTTCAGGCCCGGAGG
 CGGGGACATGAGGGATAACTGGCGCAGCGAACTGTACAAGTATAAAGTGGTCAAGATCGAGC
 CACTGGGAGTGGCACCACCCGATGCAAACGAAGAGTGGTCGGAAGGCGACGACGAAGAAGG
 GCAGTGGGCATTGGGGCCGTCTTCCTGGGGTTTCTGGGAGCAGCCGGCTCTACAATGGGAGC
 TGCAAGTATGACCCTGACAGTGACAGGCTAGGAATCTGCTGTGAGGCATCGTCCAGCAGCAGA
 GCAACCTGCTGCGAGCACCAAGACAGCAGCATCTGCTGAAGCTGACCGTGTGGGGCATC
 AAGCAGCTGCAGGCCCGGGTGCTGGCTGTGAGCGCTACCTGCGAGATCAGCAGCTGCTGGG
 GATCTGGGGATGTAGCGGCAAGCTGATTTGCTGTACAAATGTGCCTTGGAACCTCTAGTTGGA
 GCAATAGAAACCTGTCCGAAATCTGGGACAATATGACATGGCTGCAGTGGGATAAGGAGATT
 TCTAACTACACTCAGATCATCTACGGCCTGCTGGAAGAGAGTCAGAATCAGCAGGAGAAGAA
 CGAGCAGGACCTGCTGGCCCTGGAT

[00228] # Strain B.US.1998.AC10_29.AY835446 stabilized by olio6 mutations

[00229] AC10_SOSIP_olio6_mC

AVEQTWVTVYYGVVPVWKEANTTLFCASDAKAYNTEVHNVWATHACVPTDPNPQEVELENVTE
NFMWKNMVDQMEDIISLWDQSLKPCVKLTPLCVTLSCDNDVGNSTNNRWDKMEKGE
IKNCSFNITTNMRDKMQQYALFWKLDVVPPIEEGKNNNSSFTDYRLISCNTSVITQACPKVT
FEPIPIHYCAPAGFALLKCKDKKFNGTGPCKNVSTVQCTHGIKPVVSTQLLLNGSLAESEEV
IRSENFSSNARTIIVQLNTSVEIKCIA~~PN~~~~NFT~~VKGIHIGPGRAFYYMGDIIGDIRQAHCNIS
RQNWNNTLKQIAEKLREQFGNKTIVFRQSSGGDPEIVMHTFNCAGMFFYCNTAELFNSTWYA
NGTISIGGGNKTNIILPCRIKLFINMWQEVGKAMYAPPISGQIRCSSNITGLLLTRDGGGRN
QTDNQTEIFRPVGGDMKNNWRSELYKYKVVRIEPLGIAPTRCKRRVV~~GRRRRRR~~RAVGIGALS
LGFLGAAGSTMGAASMTLTVQARLLLSGIVQQQNNLLRAPEPQQHLLQDTHWGIKQLQARVL
AVEHYLRDQQLLGIWGC SGKLICCTAVPWNVSWNNRSVDDIWENMTWMQWDREISNYTSLIY
TLIEESQNQOEKNEQELLALD

[00230] GCCGTGGAGCAGACCTGGGTGACAGTGTACTATGGCGTGCCCGTGTGGAAGGAGG
CCAACACCACACTGTTCTGCGCCAGCGACGCCAAGGCCTACAACACCGAGGTGCACAACGTG
TGGGCAACCCACGCATGCGTGCTACAGATCCAAATCCCCAGGAGGTGGAGCTGGAGAACGT
GACAGAGAACCTTCAACATGTGGAAGAACAATATGGTGGACCAGATGCACGAGGACATCATCT
CTCTGTGGGACCAGAGCCTGAAGCCTTGCGTGAAGCTGACCCCACTGTGCGTGACCCCTGTCT
TGTACAGACAATGTGGGCAACGATACCAGCACAAACAATTCAGATGGGATAAGATGGAGAA
GGGCGAGATCAAGAATTGTAGCTTCAACATCACCAACAATATGAGGGACAAGATGCAGAAGC
AGTACGCCCTGTTTTGGAAGCTGGATGTGGTGCCCATCGAGGAGGGCAAGAACAATAACAGC
TCCTTCACCGACTATAGACTGATCTCTTGCAATACCAGCGTGATCACACAGGCCTGTCCAAA
GGTGACATTTGAGCCTATCCCAATCCACTACTGCGCACCAGCAGGATTTCGCACTGCTGAAGT
GTAAGGATAAGAAGTTTAATGGCACCGGCCCTTGCAAGAACGTGTCCACCGTGCACTGTACA
CACGGCATCAAGCCAGTGGTGTCTACACAGCTGCTGCTGAACGGCAGCCTGGCCGAGGAGGA
GGTGGTCATCCGGTCCGAGAATTTCTCTAATAACGCCAGAACCATCATCGTGCACTGAACA
CATCCGTGGAGATCAAGTGCATCGCCCCAATAACTTCACCGTGAAGGGCATCCACATCGGC
CCTGGCAGGGCCTTTTACTATATGGGCGACATCATCGGCGACATCAGGCAGGCCCACTGTAA
CATCTCTCGCCAGAATTGGAATAACACCCTGAAGCAGATCGCCGAGAAGCTGCGCGAGCAGT
TCGGCAATAAGACAATCGTGTTTCGGCAGTCTAGCGGCGGCGACCCAGAGATCGTGATGCAC
ACCTTCAACTGCGCCGGCATGTTCTTTTACTGTAACACAGCCGAGCTGTTTAATTCCACCTG
GTATGCCAACGGCACAATCTCTATCGGCGGCGGCAATAAGACCAACATCATCCTGCCCTGCC
GCATCAAGCTGTTTCATCAATATGTGGCAGGAAGTGGGCAAGGCAATGTACGCACCACCTATC
AGCGGACAGATCAGGTGTTCTCTAACATCACCGGCCCTGCTGCTGACACGGGACGGCGGCAG
AGGAAACCAGACCGATAATCAGACAGAGATCTTTAGACCTGTGGGCGGCGATATGAAGAATA
ACTGGCGGTCCGAGCTGTACAAGTATAAGGTGGTGAGAATCGAGCCACTGGGAATCGCACCA
ACCAGGTGCAAGAGGAGAGTGGTGGGCAGGCGCCGAGAAGGCGCGCAGTGGGCATCGGCGC
CCTGAGCCTGGGCTTTCTGGGAGCAGCAGGCAGCACAAATGGGCGCAGCCTCCATGACCCCTGA

CAGTGCAGGCCAGACTGCTGCTGTCCGGCATCGTGCAGCAGCAGAATAACCTGCTGAGGGCC
 CCCGAGCCTCAGCAGCACCTGCTGCAGGACACCCACTGGGGCATCAAGCAGCTGCAGGCCAG
 GGTGCTGGCAGTGGAGCACTATCTGCGCGATCAGCAGCTGCTGGGCATCTGGGGCTGCTCTG
 GCAAGCTGATCTGCTGTACCGCCGTGCCCTGGAACGTGTCCTGGAATAACCGCTCTGTGGAC
 GACATCTGGGAGAATATGACATGGATGCAGTGGGACCGGGAGATCAGCAACTACACCTCCCT
 GATCTATACACTGATCGAGGAGTCCCAGAACCAGCAGGAGAAGAATGAGCAGGAGCTGCTGG
 CCCTGGAT

[00231] # Resurfaced BG505 SOSIP trimer

[00232] BG505_SOSIP_MD39_SET224_wLoops_4_mC (BG505_SOSIP_R4_mC)

[00233] AENLWVTVYYGVPVWKDAETTLFCASDAKAYSTEKHNWATHACVPTDPNPQEVV
 LENVTENFNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTELNDTNTTATNSSGR
 VIEDKEIKNCSFNMTTSLRDKVQRVYSLFNKFDIVPIDNSNDSYRLISCNTSAITQACPKVS
 FEPIPIHYCAPAGFAILKCNDKEFNGTGPCKSVSTVQCTHGIRPVVSTQLLLNGSLAEFEVI
 IRSENFTNNAKTILVQLNEPVVINCTRPNNNTVKSIRIGPGQAFYYTGEIIGDIRQAHCTVS
 RETWNKTLGRVVEQLREQFRNKTIIVFNQSSGGDPEIVMHSFNCGGEFFYCNSTQLFNSTWY
 GNETETGGTNDTIGNITLPCRICKQIINMWQEVGKAMYAPPIRGQISCSNITGLILTRDGGN
 NNETNTTETFRPGGDMRDNRSELYKYKVVKIEPLGVAPTCKKRRVVGRRRRRRRAVGIGAM
 SLGFLGAAGSTMGAASLTTLTVQARNLLSGIVQQSNLLRAPEPQQHLLKDTHWGIKQLQARV
 LAVEHYLRDQQLLGIWGC SGKLI CCTNPWNTSWSNKS LDQ IWDNMTWLEWDREISNYTQLI
 YNLLEESQNQQEKNEQDLLALD

[00234] GCCGAGAATCTGTGGGTGACCGTGACTATGGCGTGCCCGTGTGGAAGGACGCAG
 AGACCACACTGTTCTGCGCCTCCGATGCCAAGGCCTACTCTACAGAGAAGCACAACGTGTGG
 GCAACCCACGCATGCGTGCTACAGACCCAAATCCCCAGGAGGTGGTGCTGGAGAACGTGAC
 CGAGAACTTTAATATGTGGAAGAACAATATGGTGAGCAGATGCACGAGGACATCATCTCTC
 TGTGGGATCAGAGCCTGAAGCCTTGCGTGAACTGACCCCACTGTGCGTGACCCTGAATTGT
 ACAGAGCTGAACGACACCAATACCACAGCCACAAACAGCTCCGGCAGAGTGATCGAGGATAA
 GGAGATCAAGAACTGTAGCTTCAATATGACCACATCCCTGCGGGACAAGGTGCAGAGAGTGT
 ACTCCCTGTTCAATAAGTTTGATATCGTGCTATCGACAACCTCCAATGATTCTTATAGACTG
 ATCAGCTGCAACACCTCCGCCATCACACAGGCCTGTCCAAAGGTGTCTTTTGAGCCTATCCC
 AATCCACTACTGCGCACCAGCAGGATTGCAATCCTGAAGTGTAACGACAAGGAGTTTAATG
 GCACCGGCCCTTGCAAGAGCGTGTCCACCGTGCACTGTACACACGGCATCCGGCCAGTGGTG
 TCTACACAGCTGCTGCTGAATGGCAGCCTGGCCGAGGAGGAAGTGATCATCAGAAGCGAGAA
 CTTCACCAACAATGCCAAGACAATCCTGGTGAGCTGAACGAGCCCGTGGTCATCAACTGCA
 CCAGGCCTAACAATAACACAGTGAAGTCTATCCGCATCGGCCAGGCCAGGCCTTTTACTAT
 ACCGGCGAGATCATCGGCGATATCAGGCAGGCCCACTGTACAGTGAGCCGCGAGACCTGGAA
 CAAGACACTGGGAAGGGTGGTGAGCAGCTGAGGGAGCAGTTCAGAAATAAGACCATCATCG
 TGTTTAACCAGTCTAGCGGCGGCGACCCAGAGATCGTGATGCACTCCTTCAACTGCGGCGGC
 GAGTTCTTTTACTGTAACTCTACCCAGCTGTTTAATAGCACATGGTATGGCAACGAGACCGA

GACAGGCGGCACCAACGATACAATCGGCAATATCACCCCTGCCCTGCAGGATCAAGCAGATCA
TCAATATGTGGCAGGAAGTGGGCAAGGCAATGTACGCACCACCTATCAGGGGACAGATCAGC
TGTTCTCTAACATCACCGGCCTGATCCTGACAAGGGACGGAGGCAATAACAATGAGACCAA
TACCACAGAGACATTAGACCCCGCGGCGGCGACATGAGGGATAACTGGCGCTCCGAGCTGT
ACAAGTATAAGGTGGTGAAGATCGAGCCACTGGGAGTGGCACCAACCAAGTGCAAGAGGAGA
GTGGTGGGCGAGGCGCCGAGAAGGCGCGCAGTGGGCATCGGCGCCATGAGCCTGGGCTTTCT
GGGAGCAGCAGGATCCACAATGGGAGCAGCCTCTCTGACCCTGACAGTGCAGGCCAGGAATC
TGCTGAGCGGCATCGTGCAGCAGCAGTCCAACCTGCTGAGGGCACCAGAGCCTCAGCAGCAC
CTGCTGAAGGACACCCACTGGGGCATCAAGCAGCTGCAGGCACGGGTGCTGGCAGTGGAGCA
CTATCTGAGAGATCAGCAGCTGCTGGGAATCTGGGGATGCAGCGGCAAGCTGATCTGCTGTA
CCAACGTGCCTTGAATACATCTTGGAGCAATAAGTCCCTGGACCAGATCTGGGATAACATG
ACCTGGCTGGAGTGGGATAGAGAGATCTCCAATTACACACAGCTGATCTATAACCTGCTGGA
GGAGTCTCAGAACCAGCAGGAGAAGAATGAGCAGGACCTGCTGGCCCTGGAT

[00235] VI: Additional trimer modifications that add functionality and that can be combined with other types of modifications

[00236] # Filling conserved glycan holes on BG505: introduction of conserved glycans missing from the BG505 strain: N241 is missing and in 97% of HIV strains, N289 is missing and in 72% of HIV strains

[00237] BG505_SOSIP_MD39_congly_mC

[00238] AENLWVTVYYGVVPWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQEIH
LENVTEEFNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNVNNTDDMRGELK
NCSFNMTTELKDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVS
FEPIPIHYCAPAGFAILKCKDKKFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEFEVI
IRSENITNNAKNILVQLNTSVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNVS
KATWNETLGKVVQQLRKHFNGNTIIRFAQSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWI
SNTSVQGSNSTGSNDSITLPCRICKIINMWQRIGQAMYAPPIQGVIKRVSNITGLILTRDGG
STNSTTETFRPGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVS
LGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARVL
AVEHYLRDQQLLGIWCGSKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
GLLEESQNQQEKNEQDLLALD

[00239] GCCGAGAACCTGTGGGTGACCGTGTACTATGGCGTGCCCGTGTGGAAGGACGCAG
AGACCACACTGTTCTGCGCCAGCGATGCCAAGGCCTACGAGACAGAGAAGCACAACGTGTGG
GCAACCCACGCATGCGTGCCTACAGACCCAAACCCCGAGGAGATCCACCTGGAGAATGTGAC
CGAGGAGTTTAACATGTGGAAGAACAATATGGTGGAGCAGATGCACGAGGACATCATCTCTC
TGTGGGATCAGAGCCTGAAGCCTTGCGTGAAGCTGACCCCACTGTGCGTGACACTGCAGTGT
ACCAACGTGACAAACAATATCACCGACGATATGAGGGGCGAGCTGAAGAATTGTTCTTTCAA
CATGACCACAGAGCTGCGGGACAAGAAGCAGAAAGTGTACAGCCTGTTTTATAGACTGGATG

TGGTGCAGATCAATGAGAACCAGGGCAATAGGAGCAACAATTCCAACAAGGAGTACCGCCTG
 ATCAATTGCAACACCTCTGCCATCACACAGGCCTGTCCTAAGGTGAGCTTCGAGCCTATCCC
 AATCCACTATTGCGCACCAGCAGGATTTCGAATCCTGAAGTGTAAGGATAAGAAGTTTAATG
 GCACCGGCCCTTGCCAGAACGTGTCCACCGTGCAGTGTACACACGGCATCAAGCCAGTGGTG
 AGCACACAGCTGCTGCTGAATGGCTCCCTGGCCGAGGAGGAAGTGATCATCCGGTCTGAGAA
 CATCACCAACAATGCCAAGAATATCCTGGTGCAGCTGAACACAAGCGTGCAGATCAATTGCA
 CCAGGCCCAACAATAACACAGTGAAGTCCATCCGCATCGGCCCTGGCCAGGCCTTTTACTAT
 ACCGGCGACATCATCGGCGACATCAGACAGGCCCACTGTAACGTGAGCAAGGCCACCTGGAA
 CGAGACACTGGGCAAGGTGGTGAAGCAGCTGCGGAAGCACTTCGGCAATAACACCATCATCA
 GATTTGCACAGAGCAGCGGAGGCGACCTGGAGGTGACCACACACTCCTTCAATTGCGGCGGC
 GAGTTCTTTTACTGTAACACATCCGGCCTGTTTAATTCTACCTGGATCTCTAACACAAGCGT
 GCAGGGCTCCAATTCTACCGGCTCCAACGATTCTATCACACTGCCATGCAGGATCAAGCAGA
 TCATCAACATGTGGCAGAGGATCGGACAGGCAATGTATGCACCACCTATCCAGGGCGTGATC
 AGATGCGTGAGCAATATCACCGGCCTGATCCTGACACGCGACGGAGGCAGCACCAACTCCAC
 CACAGAGACATTAGAGCCCCGGCGGAGGCGACATGAGGGATAACTGGAGATCCGAGCTGTACA
 AGTATAAGGTGGTGAAGATCGAGCCACTGGGAGTGGCACCAACCAGGTGCAAGAGGAGAGTG
 GTGGGCAGGCGCCGGAGAAGGCGCGCAGTGGGCATCGGAGCCGTGAGCCTGGGCTTTCTGGG
 AGCAGCAGGCTCTACAATGGGCGCAGCCAGCATGACCCTGACAGTGCAGGCCCGGAATCTGC
 TGTCCGGCATCGTGCAGCAGCAGTCTAACCTGCTGAGAGCCCCCGAGCCTCAGCAGCACCTG
 CTGAAGGACACCCACTGGGGCATCAAGCAGCTGCAGGCCAGAGTGCTGGCCGTGGAGCACTA
 CCTGAGAGATCAGCAGCTGCTGGGCATCTGGGGATGCTCCGGCAAGCTGATCTGCTGTACCA
 ATGTGCCTTGGAATCTAGCTGGAGCAATAGAAACCTGTCCGAGATCTGGGACAATATGACC
 TGGCTGCAGTGGGATAAGGAGATCTCCAACCTACACACAGATCATCTATGGCCTGCTGGAGGA
 GTCTCAGAATCAGCAGGAGAAGAACGAGCAGGACCTGCTGGCCCTGGAT

[00240] # C-terminal cysteine constructed used for conjugations. This modification can be added to any trimer.

[00241] BG505_SOSIP_D664_MD39_CtCys_mC

[00242] AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQEIH
 LENVTEEFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELK
 NCSFNMTELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVS
 FEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEVEVI
 IRSENITNNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNVS
 KATWNETLGKVVKQLRKHFNNNTIIRFAQSSGGDLEVTHSFNCGGEFFYCNTSGLFNSTWI
 SNTSVQGSNSTGSNDSITLPCRICKIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGG
 STNSTTETFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVS
 LGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARVL
 AVEHYLRDQQLLGIWCGSKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
 GLLEESQNQQEKNEQDLLALDGTKHHHHHHHC

[00243] GCCGAAAATCTGTGGGTGACTGTCTACTATGGCGTGCCTGTCTGGAAGGACGCCG
AGACCACACTGTTCTGTGCTTCCGATGCTAAGGCATACGAAACCGAGAAACACAACGTGTGG
GCAACCCATGCCTGCGTCCCAACAGACCCAAACCCCCAGGAAATCCACCTGGAGAATGTGAC
CGAGGAGTTCAACATGTGGAAGAACAATATGGTGGAAACAGATGCATGAGGACATCATTTCCC
TGTGGGATCAGTCTCTGAAGCCTTGCGTGAAACTGACCCCACTGTGCGTGACACTGCAGTGT
ACAAACGTCACCTAACAATATCACCGACGATATGCGGGGCGAACTGAAGAATTGTTCTTTCAA
CATGACTACCGAGCTGAGGGACAAGAAACAGAAAGTGACAGTCTGTTTTATCGCCTGGATG
TGGTCCAGATCAATGAAAACAGGGGAATCGAAGTAAACAATTCAAACAAGGAGTACCGGCTG
ATTAATTGCAACACCAGTGCCATCACACAGGCTTGTCCAAAAGTGTCAATTCGAGCCTATCCC
AATTCATTATTGCGCCCCCGCTGGCTTCGCCATCCTGAAGTGTAAGATAAGAAGTTCAACG
GCACCGGGCCATGCCCTTCAGTGAGCACTGTCCAGTGTACCCACGGAATTAAGCCTGTGGTC
TCCACACAGCTGCTGCTGAATGGCTCTCTGGCTGAGGAAGAAGTGATCATTAGGTCTGAGAA
CATCACTAACAATGCAAAGAATATTCTGGTGCAGCTGAACACCCCCGTCCAGATCAATTGCA
CTCGCCCTAACAATAACACCGTGAAATCTATCCGAATTGGACCCGGCCAGGCTTTTTATTAC
ACCGGCGACATCATTGGCGACATCAGACAGGCACACTGCAATGTGAGCAAGGCCACATGGAA
CGAGACTCTGGGGAAGGTGGTCAAACAGCTGCGCAAACATTTTCGGAATAACACAATCATTC
GATTTGCACAGAGCAGCGGAGGGGACCTGGAAGTGACAACTCACAGCTTCAATTGCGGAGGC
GAGTTCTTTTACTGTAACACTAGTGGCCTGTTTAATTCAACTTGGATCAGCAACACCTCCGT
GCAGGGCAGCAACAGCACCGGCTCTAACGATAGTATCACACTGCCATGTCTGGATTAAGCAGA
TCATTAACATGTGGCAGAGAATCGGGCAGGCCATGTATGCACCCCCTATCCAGGGAGTGATT
CGATGCGTGAGCAATATCACAGGCCTGATTCTGACTAGAGACGGGGGATCAACAAACAGCAC
CACAGAGACTTTCCCGCCCCGGCGGAGGAGACATGCGAGATAACTGGAGATCCGAACTGTACA
AGTATAAAGTGGTCAAGATCGAGCCACTGGGCGTGGCTCCCACCAGATGCAAACGAAGAGTG
GTCGGGAGGCGCCGACGGAGAAGGGCTGTGGGGATTGGAGCAGTCAGCCTGGGCTTTCTGGG
GGCCGCTGGATCTACAATGGGGGCAGCCAGTATGACTCTGACCGTGCAGGCCAGGAATCTGC
TGTCAGGAATCGTGCAGCAGCAGAGCAACCTGCTGAGGGCTCCCGAACCCACAGCAGCACCTG
CTGAAGGACACCCACTGGGGCATCAAGCAGCTGCAGGCAAGAGTGCTGGCCGTCGAGCATTA
CCTGAGGGATCAGCAGCTGCTGGGCATCTGGGGATGCAGCGGAAAGCTGATTTGCTGTACAA
ATGTGCCTTGGAATCTAGTTGGAGCAATCGCAACCTGTCCGAAATCTGGGACAATATGACA
TGGCTGCAGTGGGATAAGGAGATTAGCAACTACACTCAGATCATCTACGGCCTGCTGGAAGA
GTCCCAGAATCAGCAGGAGAAGAACGAGCAGGACCTGCTGGCCCTGGATGGTACCAAGCACC
ACCATCACCATCACTGT

[00244] # Computational designed mutation knocks out binding of CD4 receptor, yet retains antigenic profile of BG505 CD4bs. In this case Applicants illustrate the mutation on the MD39 background.

[00245] BG505_SOSIP_D664_MD39_CD4KO4_mC

[00246] AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHL
LENVTEEFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELK

NCSFNMTTELKQKQVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVS
FEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEFEVI
IRSENITNNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNVS
KATWNETLGKVVKQLRKHFNGNTIIRFAQSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWI
SNTSVQGSNSTGSNDSITLPCRICKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGG
STNSTTETFRPGGTMDRDNWRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVS
LGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARVL
AVEHYLRDQQLLGIWGC SGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
GLLEESQNQQEKNEQDLLALD

[00247] GCCGAAAATCTGTGGGTGACTGTCTACTATGGCGTGCCTGTCTGGAAGGACGCCG
AGACCACACTGTTCTGTGCTTCCGATGCTAAGGCATACGAAACCGAGAAACACAACGTGTGG
GCAACCCATGCCTGCGTCCCAACAGACCCAAACCCCGAGGAAATCCACCTGGAGAATGTGAC
CGAGGAGTTCAACATGTGGAAGAACAATATGGTGGAACAGATGCATGAGGACATCATTTCCC
TGTGGGATCAGTCTCTGAAGCCTTGCCTGAACTGACCCCACTGTGCGTGACACTGCAGTGT
ACAAACGTCACTAACAATATCACCGACGATATGCGGGGCGAACTGAAGAATTGTTCTTTCAA
CATGACTACCGAGCTGAGGGACAAGAAACAGAAAGTGACAGTCTGTTTTATCGCCTGGATG
TGGTCCAGATCAATGAAAACAGGGGAATCGAAGTAACAATTCAAACAAGGAGTACCGGCTG
ATTAATTGCAACACCAGTGCCATCACACAGGCTTGTCCAAAAGTGTCAATTCGAGCCTATCCC
AATTCATTATTGCGCCCCCGCTGGCTTCGCCATCCTGAAGTGTAAGATAAGAAGTTCAACG
GCACCGGGCCATGCCCTTCAGTGAGCACTGTCCAGTGTAACCCACGGAATTAAGCCTGTGGTC
TCCACACAGCTGCTGCTGAATGGCTCTCTGGCTGAGGAAGAAGTGATCATTAGGTCTGAGAA
CATCACTAACAATGCAAAGAATATTCTGGTGAGCTGAACACCCCGTCCAGATCAATTGCA
CTCGCCCTAACAATAACACCGTGAAATCTATCCGAATTGGACCCGGCCAGGCTTTTTATTAC
ACCGGCGACATCATTGGCGACATCAGACAGGCACACTGCAATGTGAGCAAGGCCACATGGAA
CGAGACTCTGGGGAAGGTGGTCAAACAGCTGCGCAAAATTTGGAATAACACAATCATTC
GATTTGCACAGAGCAGCGGAGGGGACCTGGAAGTGACAACTCACAGCTTCAATTGCGGAGGC
GAGTTCTTTTACTGTAACACTAGTGGCCTGTTTAATTCAACTTGGATCAGCAACACCTCCGT
GCAGGGCAGCAACAGCACCGGCTCTAACGATAGTATCACACTGCCATGTCCGATTAAGCAGA
TCATTAACATGTGGCAGAGAATCGGGCAGGCCATGTATGCACCCCTATCCAGGGAGTGATT
CGATGCGTGAGCAATATCACAGGCCTGATTCTGACTAGAGACGGGGGATCAACAAACAGCAC
CACAGAGACTTTCCGGCCCCGGCGGAACCGACATGCGAGATAACTGGAGATCCGAACTGTACA
AGTATAAAGTGGTCAAGATCGAGCCACTGGGCGTGGCTCCACCAGATGCAAACGAAGAGTG
GTCGGGAGGCGCCGACGGAGAAGGGCTGTGGGGATTGGAGCAGTCAGCCTGGGCTTTCTGGG
GGCCGCTGGATCTACAATGGGGGACGCCAGTATGACTCTGACCGTGCAGGCCAGGAATCTGC
TGTCAGGAATCGTGCAGCAGCAGAGCAACCTGCTGAGGGCTCCCGAACCCCGCAGCAGCCTG
CTGAAGGACACCCACTGGGGCATCAAGCAGCTGCAGGCAAGAGTGCTGGCCGTCGAGCATTA
CCTGAGGGATCAGCAGCTGCTGGGCATCTGGGGATGCAGCGGAAAGCTGATTTGCTGTACAA
ATGTGCCTTGGAACCTAGTTGGAGCAATCGCAACCTGTCCGAAATCTGGGACAATATGACA

TGGCTGCAGTGGGATAAGGAGATTAGCAACTACACTCAGATCATCTACGGCCTGCTGGAAGA
GTCCCAGAATCAGCAGGAGAAGAACGAGCAGGACCTGCTGGCCCTGGAT

[00248] # Computational designed mutation knocks out binding of CD4 receptor, yet retains antigenic profile of BG505 CD4bs. Here the mutation is shown on the oio6 background.

[00249] BG505_SOSIP_D664_oio6_CD4KO4_mC

[00250] AENLWTVYYGVPVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQEIHL
LENVTEEFNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELK
NCSFNMTTELKDKKQKVYSLFWRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVS
FEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEFEVI
IRSENITNNAKNILVQLNTPVQINCTAPNNFTVKSIRIGPGQAFYYMGDIIGDIRQAHCNVS
KATWNETLGKVVQLRKHFGNNTIIRFAQSSGGDLEVTTTHSFNCGGMFFYCNTSGLFNSTWI
SNTSVQGSNSTGSNDSITLPCRIKLIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGG
STNSTTETFRPGGTDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVS
LGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARVL
AVEHYLRDQQLLGIWGC SGKLI CCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
GLLEESQNQQEKNEQDLLALD

[00251] GCCGAAAACCTGTGGGTGACTGTCTACTATGGCGTGCCCGTCTGGAAGGACGCAG
AGACCACACTGTTCTGCGCTAGCGATGCTAAGGCCTACGAGACCGAGAAACACAACGTGTGG
GCAACCCATGCCTGCGTGCCTACAGACCCAAATCCCCAGGAAATCCACCTGGAGAACGTGAC
CGAGGAGTTCAACATGTGGAAGAACAATATGGTGGAACAGATGCATGAGGACATCATTTCCC
TGTGGGATCAGTCTCTGAAGCCCTGCGTGAAACTGACCCCTCTGTGCGTCACACTGCAGTGT
ACAAACGTGACTAACAATATCACCGACGATATGCGGGGCGAACTGAAGAACTGTTCTTTCAA
TATGACTACCGAGCTGCGCGACAAGAAACAGAAAGTGACAGTCTGTTTTGGCGACTGGATG
TGGTCCAGATCAACGAAAATCAGGGGAACCGGAGTAACAACCTCAAATAAGGAGTATAGACTG
ATCAACTGCAATACCAGTGCCATTACACAGGCTTGTCCTAAAGTGTCATTGAGCCTATCCC
AATTCACTACTGCGCCCCAGCTGGCTTCGCCATCCTGAAGTGTAAGATAAGAAGTTCAACG
GCACCGGGCCATGCCCTTCAGTGAGCACTGTCCAGTGTACCCACGGAATTAAGCCTGTGGTC
TCCACACAGCTGCTGCTGAACGGCTCTCTGGCTGAGGAAGAAGTGATCATTAGATCCGAGAA
TATCACTAACAATGCAAAGAACATTCTGGTGAGCTGAATACCCAGTCCAGATCAACTGCA
CTGCCCCCAACAATTTACCGTGAAATCTATCCGGATTGGACCAGGCCAGGCTTTTTACTAT
ATGGGCGACATCATTTGGGGATATTAGACAGGCACACTGTAACGTGAGCAAGGCCACATGGAA
TGAAACTCTGGGGAAGGTGGTCAAACAGCTGCGAAAACATTTGCGAAACAATACAATCATTC
GATTTGCACAGAGCAGCGGAGGGGACCTGGAGGTGACAACTCACAGCTTCAACTGCGGAGGC
ATGTTCTTTTATTGTAATACTAGTGGCCTGTTTAACTCAACTTGGATCAGCAATACCTCCGT
GCAGGGCAGCAACAGCACCGGCTCTAATGATAGTATCACACTGCCATGCAGAATTAAGCTGA
TCATTAATATGTGGCAGAGGATCGGGCAGGCTATGTACGCACCCCTATCCAGGGAGTGATT
CGGTGCGTGAGCAACATCACAGGCCTGATTCTGACTAGAGACGGGGGATCAACAAATAGCAC

CACAGAGACTTTTCAGGCCCGGCGGAACCGACATGCGAGATAACTGGCGATCCGAACTGTACA
 AGTATAAAGTGGTCAAGATCGAGCCTCTGGGAGTGGCACCAACCCGATGCAAACGAAGAGTG
 GTCGGGAGGCGCCGACGGAGAAGGGCTGTGGGGATTGGAGCAGTCTCTCTGGGCTTTCTGGG
 GGCCGCTGGATCTACAATGGGGGCAGCCAGTATGACTCTGACCGTCCAGGCAAGGAACCTGC
 TGTCAGGAATCGTGCAGCAGCAGAGCAATCTGCTGCGCGCTCCAGAACCCGAGCAGCACCTG
 CTGAAGGACACCCATTGGGGCATCAAGCAGCTGCAGGCAAGGGTGCTGGCAGTCGAGCACTA
 CCTGCGAGATCAGCAGCTGCTGGGCATCTGGGGATGCAGCGGAAAGCTGATTTGCTGTACAA
 ACGTGCCCTGGAACAGCAGCTGGAGCAACCGGAATCTGTCCGAAATCTGGGACAACATGACA
 TGGCTGCAGTGGGATAAGGAGATTAGCAATTACACTCAGATCATCTACGGCCTGCTGGAAGA
 GTCCCAGAACCAGCAGGAGAAGAACGAGCAGGACCTGCTGGCCCTGGAT

[00252] # Filling conserved glycan holes on AC10: introduction of conserved glycans missing from the AC10 strain: N295 is missing and in 62% of HIV strains, N386 is missing and in 86% of HIV strains

[00253] AC10_SOSIP_olio6_congly_mC

[00254] AVEQTWVTVYYGVPVWKEANTTLFCASDAKAYNTEVHNVWATHACVPTDPNPQEV
 ELENVTENFNMWKNMVDQMHEDIISLWDQSLKPCVKLTPLCVTLSCVDNVGNDTSTNNSRW
 DKMEKGEIKNCSFNITTNMRDKMQKQYALFWKLDVVPPIEEGKNNNSSFTDYRLISCNSTSVIT
 QACPKVTFEPIPIHYCAPAGFALLKCKDKKFNGTGPCKNVSTVQCTHGIKPVVSTQLLLNGS
 LAEEEVVIRSENFSSNNARTIIVQLNTSVEINCTAPNNFTVKGIHIGPGRAFYYMGDIIGDIR
 QAHCNISRQNWNNNTLKQIAEKLREQFGNKTIIVFRQSSGGDPEIVMHTFNCAGMFFYCNTSEL
 FNSTWYANGTISIGGGNKNTNIIILPCRIKLFINMWQEVGKAMYAPPISGQIRCSSNITGLLLT
 RDGGRGNQTDNQTIEIFRPVGGDMKNNWRSELYKYKVVRIEPLGIAPTRCKRRVVGRRRRRRA
 VGIGALSLGFLGAAGSTMGAASMTLTQARLLLSGIVQQQNNLLRAPEPQQHLLQDTHWGIK
 QLQARVLAVEHYLRDQQLLGIWCGSKLICCTAVPWNVSWNNRSVDDIWENMTWMQWDREIS
 NYTSLIYTLIEESQNQQEKNEQELLALD

[00255] GCCGTGGAGCAGACCTGGGTGACAGTGTACTATGGCGTGCCCGTGTGGAAGGAGG
 CCAACACCACACTGTTCTGCGCCAGCGACGCCAAGGCCTACAACACCGAGGTGCACAACGTG
 TGGGCAACCCACGCATGCGTGCTACAGATCCAAATCCCAGGAGGTGGAGCTGGAGAACGT
 GACAGAGAACTTCAACATGTGGAAGAACAAATATGGTGGACCAGATGCACGAGGACATCATCT
 CTCTGTGGGACCAGAGCCTGAAGCCTTGCGTGAAGCTGACCCCACTGTGCGTGACCCTGTCT
 TGTACAGACAATGTGGGCAACGATAACCAGCACAAACAATTCCAGATGGGATAAGATGGAGAA
 GGGCGAGATCAAGAATTGTAGCTTCAACATCACCACAAATATGAGGGACAAGATGCAGAAGC
 AGTACGCCCTGTTTTGGAAGCTGGATGTGGTGCCCATCGAGGAGGGCAAGAACAATAACAGC
 TCCTTCACCGACTATAGACTGATCTCTTGCAATACCAGCGTGATCACACAGGCCTGTCCAAA
 GGTGACATTTGAGCCTATCCCAATCCACTACTGCGCACCAGCAGGATTTCGCACTGCTGAAGT
 GTAAGGATAAGAAGTTTAATGGCACCGGCCCTTGCAAGAACGTGTCCACCGTGCAGTGTACA
 CACGGCATCAAGCCAGTGGTGTCTACACAGCTGCTGCTGAACGGCAGCCTGGCCGAGGAGGA
 GGTGGTCATCCGGTCCGAGAATTTCTCTAATAACGCCAGAACCATCATCGTGCAGCTGAACA

CATCCGTGGAGATCAACTGCACCGCCCCCAATAACTTCACCGTGAAGGGCATCCACATCGGC
 CCTGGCAGGGCCTTTTACTATATGGGCGACATCATCGGCGACATCAGGCAGGCCCCACTGTAA
 CATCTCTCGCCAGAATTGGAATAACACCCTGAAGCAGATCGCCGAGAAGCTGCGCGAGCAGT
 TCGGCAATAAGACAATCGTGTTTCGGCAGTCTAGCGGCGGCGACCCAGAGATCGTGATGCAC
 ACCTTCAACTGCGCCGGCATGTTCTTTTACTGTAAACAAGCGAGCTGTTTAATTCCACCTG
 GTATGCCAACGGCACAATCTCTATCGGCGGCGGCAATAAGACCAACATCATCCTGCCCTGCC
 GCATCAAGCTGTTTCATCAATATGTGGCAGGAAGTGGGCAAGGCAATGTACGCACCACCTATC
 AGCGGACAGATCAGGTGTTCTCTAACATCACCGGCCCTGCTGCTGACACGGGACGGCGGCAG
 AGGAAACCAGACCGATAATCAGACAGAGATCTTTAGACCTGTGGGCGGCGATATGAAGAATA
 ACTGGCGGTCCGAGCTGTACAAGTATAAGGTGGTGAGAATCGAGCCACTGGGAATCGCACCA
 ACCAGGTGCAAGAGGAGAGTGGTGGGCGAGGCGCCGGAGAAGGCGCGCAGTGGGCATCGGCGC
 CCTGAGCCTGGGCTTTCTGGGAGCAGCAGGCAGCACAATGGGCGCAGCCTCCATGACCCTGA
 CAGTGCAGGCCAGACTGCTGCTGTCCGGCATCGTGCAGCAGCAGAATAACCTGCTGAGGGCC
 CCCGAGCCTCAGCAGCACCTGCTGCAGGACACCCACTGGGGCATCAAGCAGCTGCAGGCCAG
 GGTGCTGGCAGTGGAGCACTATCTGCGCGATCAGCAGCTGCTGGGCATCTGGGGCTGCTCTG
 GCAAGCTGATCTGCTGTACCGCCGTGCCCTGGAACGTGTCCTGGAATAACCGCTCTGTGGAC
 GACATCTGGGAGAATATGACATGGATGCAGTGGGACCGGGAGATCAGCAACTACACCTCCCT
 GATCTATACACTGATCGAGGAGTCCCAGAACCAGCAGGAGAAGAATGAGCAGGAGCTGCTGG
 CCCTGGAT

[00256] Type VII sequences

[00257] 191084_SOSIP_MD39

TENLWVTVYYGVPVWRDAETTLFCASDAKAYDTEMHNVWATHACVPTDPNPQEIDLENVTEK
 FNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTAITNDTRGNETGINRTVETTEM
 TNCFSNMTELDRKKKVNALFYKLDIVQIGENSSSQYRLINCNTSVITQACPKVTFEPIPI
 HYCAPAGFAILKCKDKEFNGTGTCRNVSSVQCTHGIKPVVSTQLLLNGSLAEGQVIIRSENI
 SDNAKTIIVQLNESVPINCTRPNNNTVRGIHLGPGQTFYTDIIGDIRQAHCNVSESKWNKA
 LQEVVKQLRQHWNKTIIFKSSSGDLEITTHSFNCGGEFFYCNTSGLFNSTWNIAGNRNDT
 KSNETITLPCRIKQIVNVWQRVGQAIYAPPIAGVIRCNSNITGLLLVRDGGATNNTDETRP
 GGNMRDNWRSELYKYKVVKIEPLGVAPTRCRRRVVERRRRRAVGLGAVSIGFLGAAGSTM
 GAASVTLTVQARQLLSGIVQQSNLLRAPEPQQHLLKDTHWGIKQLQARVLAVEHYLKDQQL
 LGIWGCSGLICCTTPWNSSWSNKSQNEIWDNMTWLQWDKEISNYTQLIYSLIEESQNQQE
 KNEQELLALD

[00258] 001428_SOSIP_MD39

[00259] VENLWVTVYYGVPVWKEARTTLFCASDAKAYETEVHNVWATHACVPTDPNPQEMV
 LGNVTENFNMWKNDMVDQMHEDEVISLWAQSLKPCVKLTPLCVTLECTQVNATQGNNTQVNVT
 QVNGDEMKNCSFNNTTTEIRDKKQKAYALFYRLDLVPLERENRGDSNSASKYILINCNTSAIT
 QACPKVNFDPPIPIHYCTPAGYAILKCNKTFNGTGSCNNVSTVQCTHGIKPVVSTQLLLNGS
 LAEEEEIIIRSENLTDNVKTIIVHLDQSVEIVCTRPNNNTVKSIRIGPGQTFYYTGDIIGNIR

EAHCNISEKKWHEMLRRVSEKLAEHFPNKTIKFTSSSGDLEITTHSFNCRGEFFYCNTSGL
FNSTYMPNGTYMPNGTNNSNSTIILPCRIKQIINMWQEVGRAMYAPPIAGNITCNSNITGLL
LVRDGGKNNNTEIFRPGGDMRDNRSELYKYKVVEIKPLGVAPTRCKRRVVGRRRRRRRAVG
LGAVSLGFLGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLQAPEPQQHLLQDTHWGIKQL
QTRVLAIEHYLKDQQLLGIWGCSGKLICCTAVPWNSSWSNKS LTDIWDNMTWMQWDREVSNY
TGIIYRLLEDSQNQQERNEQDLLALD

[00260] AC10_SOSIP_MD39

AVEQTWVTVYYGVPVWKEANTTLFCASDAKAYNTEVHNVWATHACVPTDPNPQEVELENTVE
NFMWKNMVDQM HEDIISLWDQSLKPCVKLTPLCVTLSCDNDVGNSTNNSRWDKMEKGE
IKNCSFNITTNMRDKMQKYALFYKLDVVPPIEEGKNNNSSFTDYRLISCNTSVITQACPKVT
FEPIPIHYCAPAGFALLKCKDKKFNGTGPKNVSTVQCTHGIKPVVSTQLLNGSLAEEEVV
IRSENFSSNARTIIVQLNTSVEIKCIRPNNTTVKGIHIGPGRAFYYTGDIIGDIRQAHCNIS
RQNWNTLTKQIAEKLREQFGNKTIVFRQSSGGDPEIVMHTFNCAGEFFYCNTAELFNSTWYA
NGTISIGGGNKTNIILPCRIKQFINMWQEVGKAMYAPPISGQIRCSSNITGLLLTRDGGRGN
QTDNQTEIFRPVGGDMKNNWRSELYKYKVVRIEPLGIAPTRCKRRVVGRRRRRRRAVGIGALS
LGFLGAAGSTMGAASMTLTVQARLLLSGIVQQQNNLLRAPEPQQHLLQDTHWGIKQLQARVL
AVEHYLRDQQLLGIWGCSGKLICCTAVPWNVSWNNRSVDDIWENMTWMQWDREISNYTSLIY
TLIEESQNQQEKNEQELLALD

[00261] AC10_SOSIP_olio6

[00262] AVEQTWVTVYYGVPVWKEANTTLFCASDAKAYNTEVHNVWATHACVPTDPNPQEV
ELENVTENFMWKNMVDQM HEDIISLWDQSLKPCVKLTPLCVTLSCDNDVGNSTNNSRW
DKMEKGEIKNCSFNITTNMRDKMQKYALFWKLDVVPPIEEGKNNNSSFTDYRLISCNTSVIT
QACPKVTFEPIPIHYCAPAGFALLKCKDKKFNGTGPKNVSTVQCTHGIKPVVSTQLLNGS
LAEEEVVIRSENFSSNARTIIVQLNTSVEIKCIAPNNFTVKGIHIGPGRAFYYMGDIIGDIR
QAHCNISRQNWNTLTKQIAEKLREQFGNKTIVFRQSSGGDPEIVMHTFNCAGMFFYCNTAEL
FNSTWYANGTISIGGGNKTNIILPCRIKLFINMWQEVGKAMYAPPISGQIRCSSNITGLLLT
RDGGRGNQTDNQTEIFRPVGGDMKNNWRSELYKYKVVRIEPLGIAPTRCKRRVVGRRRRRRRA
VGIGALS LGFLGAAGSTMGAASMTLTVQARLLLSGIVQQQNNLLRAPEPQQHLLQDTHWGIK
QLQARVLAVEHYLRDQQLLGIWGCSGKLICCTAVPWNVSWNNRSVDDIWENMTWMQWDREIS
NYTSLIYTLIEESQNQQEKNEQELLALD

[00263] ZM197M_MD39

[00264] MEQLWVTVYYGVPVWKEAKATLFCASDAKAYEKEVHNVWATHACVPTDPNPQEIP
LGNVTENFMWKNMADQM HEDIISLWDQSLKPCVKLTPLCVTLNCSDATSNNTKNATNTNT
TSTDNRNATSNDTEMKGEIKNCTFNITTEVRDRKTKQRALFYKLDVVPLEEEKNSSSKNSSY
KEYRLISCNTSTITQACPKVSFDPIPIHYCAPAGYAILKCNKTFNGTGPCNVSTVQCTHG
IKPVVSTQLLNGSLAEEEEIIIRSENLTDNKTIIIVHLNESVEINCTRPNNNTVKSVRIGPG
QTFYTGTEIIGDIRQAHCNLSKSNWTTTLKRIEKKLKEHFNNATIKFESSAGGDLEITTHSF
NCRGEFFYCNTSGLFNSSLLNDTDGTSNSTSNATITLPCRIKQIINMWQEVGRAMYASPIAG

IITCKSNITGLLLTRDGGNKSAGIETFRPGGGNMKDNWRSELYKYKVVEIKPLGIAPTSCRR
RVVERRRRRRRAAGIGAVSLGFLGAAGSTMGAASVMLTVQARQLLSGIVQQQSNLLRAPEPQQ
HMLQDTHWGIKQLQTRVLAIEHYLKDQQLLGLWGCSGKLI CCTAVPWNTSWSNKSDEIWDN
MTWMQWDREIDNYTQVIYQLLEVSQNQQEKNENDLLALD

[00265] B41_SOSIP_D664_MD39

[00266] AAKKWTVYYGVPVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNPQEIV
LGNVTENFNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCCNNVNTNNTNNTNATI
SDWEKMETGEMKNCSFNVTTTSIRDKIKKEYALFYKLDVVPLENKNNINNTNITNYRLINCNT
SVITQACPKVSFEPIPIHYCAPAGFAILKCNSTFNGSGPCTNVSTVQCTHGIRPVVSTQLL
LNGSLAEEEEIVIRSENITDNAKTIIIVQLNEAVEINCTRPNNNTVKSIIHIGPGRAFYYTGDI
GNIRQAHCNISKARWNETLGQIVAKLEEQFPNKTIIIFNHSSGGDPEIVTHSFNCGGEFFYC
TTPLFNSTWNNTRTDDYPTGGEQNITLQCRIKQIINMWQGVGKAMYAPPIRGQIRCSSNITG
LLLTRDGGRDQNGTETFRPGGGNMRDNWRSELYKYKVVKIEPLGIAPTACKRRVQRRRRRR
AVGLGAFSLGFLGAAGSTMGAASMLTVQARLLLSGIVQQQNNLLRAPEPQQHMLQDTHWGI
KQLQARVLAVEHYLRDQQLLGIWGCSGKIICCTNVPWNDSWSNKTINEIWDNMTWMQWEKEI
DNYTQHIYTLLEVSQIQQEKNQEELLELD

[00267] Type VIII sequences

[00268] #gp120-gp41 linker optimized by mammalian display directed evolution

[00269] BG505_SOSIP_MD39_link14

[00270] AENLWTVYYGVPVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQEIH
LENVTEEFNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNVNTNITDDMRGELK
NCSFNMTTEL RDKKQKVYSLFYRLDVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVS
FEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGKIPVVSTQLLLNGSLAEEEEVI
IRSENITNNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNVS
KATWNETLGKVVKQLRKHFGNNTIIRFAQSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWI
SNTSVQGSNSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGG
STNSTTETFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVG**SHSGSGSGSGSGHA**
AVGIGAVSLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGI
KQLQARVLAVEHYLRDQQLLGIWGCSGKLI CCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEI
SNYTQIIYGLLEESQNQQEKNQDLLALD

[00271] #The following "CP" designs are all circular permutations of gp140 that eliminate the need for a cleavage site while maintaining native-like structure.

[00272] #Circular permutation 1, variant 1

[00273] BG505_SOSIP_MD39_CP1.1

[00274] VSLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGI
IKQLQARVLAVEHYLRDQQLLGIWGCSGKLI CCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEI
ISNYTQIIYGLLEESQNQQEKNQDLLALD**GSGSGSGSGSGSGSSA**ENLWTVYYGVPVWKD

AETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLNVTEEFNMWKNNMVEQMHEDI I
 SLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELKNCSFNMTTEL RDKKQKVYSLFYRL
 DVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDKKF
 NGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEEEVIIIRSENITNNAKNILVQLNTPVQIN
 CTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNVSKATWNETLGKVVKQLRKHFNGNTI
 IRFAQSSGGDLEVTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQGSNSTGSNDSITLPCRIK
 QIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGGGDMRDNRSEL
 YKYKVVKIEPLGVAPTRCKRRVVG

[00275] #Circular permutation 1, variant 2

[00276] BG505_SOSIP_MD39_CP1.2

[00277] VSLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWG
 IKQLQARVLAVEHYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKE
 ISNYTQIIYGLLEESQNQQEKNEQDLGSGSGSGSGSGSGSGSGSLWVTVYYGVPVWKDAETT
 LFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLNVTEEFNMWKNNMVEQMHEDI I SLWD
 QSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELKNCSFNMTTEL RDKKQKVYSLFYRLDVVQ
 INENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDKKFNGTG
 PCPSVSTVQCTHGIKPVVSTQLLLNGSLAEEEVIIIRSENITNNAKNILVQLNTPVQINCTRP
 NNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNVSKATWNETLGKVVKQLRKHFNGNTIIRFA
 QSSGGDLEVTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQGSNSTGSNDSITLPCRIKQIIN
 MWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGGGDMRDNRSELYKYK
 VKIEPLGVAPTRCKRR

[00278] # Circular permutation 2

[00279] BG505_SOSIP_MD39_CP2

[00280] VSLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWG
 IKQLQARVLAVEHYLRDQQLLGIWGCSGKLICCTGSGSVTVYYGVPVWKDAETTLFCASDAK
 AYETEKHNVWATHACVPTDPNPQEIHLNVTEEFNMWKNNMVEQMHEDI I SLWDQSLKPCVK
 LTPLCVTLQCTNVTNNITDDMRGELKNCSFNMTTEL RDKKQKVYSLFYRLDVVQINENQGNR
 SNNSNKEYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTV
 QCTHGIKPVVSTQLLLNGSLAEEEVIIIRSENITNNAKNILVQLNTPVQINCTRPNNNTVKS
 IRIGPGQAFYYTGDIIGDIRQAHCNVSKATWNETLGKVVKQLRKHFNGNTIIRFAQSSGGDLE
 VTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQGSNSTGSNDSITLPCRIKQIINMWQRIGQA
 MYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGGGDMRDNRSELYKYKVVKIEPLG
 VAPTRCKRGSGSGVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEK
 NEQDLLALD

[00281] # Circular permutation 3

[00282] BG505_SOSIP_MD39_CP3

[00283] VSLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQSNLLRAPEPQQHLLKDTHWG
IKQLQARVLAVEHYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKE
ISNYTQIIYGLLEESGGSGSGAGGLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNWAT
HACVPTDPNPQEIHLNVTEEFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTN
VTNNITDDMRGELKNCSFNMTTEL RDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLIN
CNTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVST
QLLLNGSLAEEVVIIRSENITNNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPGQAFYYTG
DIIGDIRQAHCNVSKATWNETLGKVVKQLRKHFNGNTIIRFAQSSGGDLEVTTHSFNCGGEF
FYCNTSGLFNSTWISNTSVQGSNSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQGVIRC
VSNITGLILTRDGGSTNSTTETFRPGGGDMRDNRSELYKYKVKIEPLGVAPTRCKRRVVG
GGSGSGQNQQEKNEQDLLALD

[00284] Type IX sequences

[00285] # BG505 MD39 GRSF4 adds glycosylation sites at positions 80, 241, 289, 657, and 665 relative to MD39.

[00286] BG505_MD39_GRSF4

[00287] AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNSSE
IHLNVTEEFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRG
ELKNCSFNMTTEL RDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQAC
PKVSFEPIPIHYCAPAGFAILKCKDKKFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLA
EEVVIIRSENITNNAKNILVQLNTSVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQ
AHCNVSKATWNETLGKVVKQLRKHFNGNTIIRFAQSSGGDLEVTTHSFNCGGEFFYCNTSG
LFNSTWISNTSVQGSNSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQGVIRC
VSNITGLILTRDGGSTNSTTETFRPGGGDMRDNRSELYKYKVKIEPLGVAPTRCKRRVVGRRRRR
RAVGIGAVSLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQSNLLRAPEPQQHLLKDTHW
GIKQLQARVLAVEHYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWD
KEISNYTQIIYGLLEESQNQQEKNQSSLLALDNGS

[00288] # BG505 MD39 GRSF7 adds one additional glycosylation site, at position 634, on top of MD39 GRSF4. Adding this glycosylation site is accomplished by the mutation E634N, but that breaks a salt bridge with R617, so we added the R617A mutation. In total, BG505 MD39 GRSF7 contains six extra glycosylation sites relative to BG505 MD39.

[00289] BG505_MD39_GRSF7

[00290] AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNSSE
IHLNVTEEFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRG
ELKNCSFNMTTEL RDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQAC
PKVSFEPIPIHYCAPAGFAILKCKDKKFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLA
EEVVIIRSENITNNAKNILVQLNTSVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQ
AHCNVSKATWNETLGKVVKQLRKHFNGNTIIRFAQSSGGDLEVTTHSFNCGGEFFYCNTSG

LFNSTWISNTSVQGSNSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQGVIKRVSNITG
LILTRDGGSTNSTTETFRPGGGDMRDNRSELYKYKVKIEPLGVAPTRCKRRVVGRRRRR
RAVGIGAVSLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKTDTHW
GIKQLQARVLAVEHYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNANLSEIWDNMTWLQWD
KNISNYTQIIYGLLEESQNQQEKNQSLALDNGS

[00291] # BG505 CP1.2 GRSF4 adds five glycosylation sites relative to BG505 CP1.2, but some of the additional glycosylation sites are different than those in BG505_MD39_GRSF4.

[00292] BG505_MD39_CP1.2_GRSF4

[00293] VSLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKTDTH
WGIKQLQARVLAVEHYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQW
DKEISNYTQIIYGLLEESQNQQEKNQDLGGNGSGGGSGSGGNGSSGLWVTVYYGVPVWKD
AETTLFCASDAKAYETEKHNWATHACVPTDPNSSEIHLENVTEEFNMWKNMVEQMHEIDI
ISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELKNCSFNMTTEL RDKKQKVYSLFY
RLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCKD
KKFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEVVIIRSENITNNAKNILVQLNTS
VQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNVSKATWNETLGKVVQQLRKHF
GNNTIIRFAQSSGGDLEVTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQGSNSTGSNDSIT
LPCRIKQIINMWQRIGQAMYAPPIQGVIKRVSNITGLILTRDGGSTNSTTETFRPGGGDMR
DNWRSELYKYKVKIEPLGVAPTRCKRR

[00294] #Merges the BG505_MD39_link14 modifications with glycan masking modifications in BG505_MD39_GRSF4.

[00295] BG505_MD39_link14_GRSF4

[00296] AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNSSE
IHLENVTEEFNMWKNMVEQMHEIDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRG
ELKNCSFNMTTEL RDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQAC
PKVSFEPIPIHYCAPAGFAILKCKDKKFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLA
EEVVIIRSENITNNAKNILVQLNTSVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQ
AHCNVSKATWNETLGKVVQQLRKHFGNNTIIRFAQSSGGDLEVTTHSFNCGGEFFYCNTSG
LFNSTWISNTSVQGSNSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQGVIKRVSNITG
LILTRDGGSTNSTTETFRPGGGDMRDNRSELYKYKVKIEPLGVAPTRCKRRVVGSHSGS
GGSGSGGHA AVGIGAVSLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQ
HLLKTDTHWGIKQLQARVLAVEHYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWD
NMTWLQWDKEISNYTQIIYGLLEESQNQQEKNQSLALDNGS

[00297] #Merges the BG505_MD39_link14 modifications with glycan masking modifications in BG505_MD39_GRSF7.

[00298] BG505_MD39_link14_GRSF7

[00299] AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNSSE
 IHLENVTEEFNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNVNTNITDDMRG
 ELKNCSFNMTTEL RDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQAC
 PKVSFEPIPIHYCAPAGFAILKCKDKKFNGTGPCQNVSTVQCTHGIKPVVSTQLLNGSLA
 EEEVIIRSENITNNAKNILVQLNTSVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQ
 AHCNVSKATWNETLGKVVQQLRKHFNGNTIIRFAQSSGGDLEVTTHSFNCGGEFFYCNTSG
 LFNSTWISNTSVQGSNSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQGVI RCVSNITG
 LILTRDGGSTNSTTETFRPGGDMRDNRSELYKYKVKIEPLGVAPTRCKRRVVGSHSGS
 GSGSGGHA AVGIGAVSLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQ
 HLLK DTHWGIKQLQARVLAVEHYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNANLSEIWD
 NMTWLQWDKNISNYTQIIYGLLEESQNQQEKNNQSLALDNGS

[00300] Type X sequences

[00301] #Cleavage-indepdent MD39 with "quiet" loops.

[00302] # V1: D141N, R143S (as in VLC1-03), V1 is already short this adds 1 extra
 glycan to c-term of loop

[00303] # V2b: Use VLC3-13 (hotspot, single glycan and 8 AA shorter)

[00304] # V4: Use VLC3-13 (hotspot, 2 glycans and 5 AA shorter)

[00305] BG505_MD39_CP1.2_GRSF4_qLoops1

[00306] VSLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLK DTHWG
 IKQLQARVLAVEHYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKE
 ISNYTQIIYGLLEESQNQQEKNEQDLGGNGSGGGSGSGNGSSGLWVTVYYGVPVWKDAETT
 LFCASDAKAYETEKHNVWATHACVPTDPNSSEIHLENVTEEFNMWKNMVEQMHEDIISLWD
 QSLKPCVKLTPLCVTLQCTNVNTNITDNMTGELKNCSFNMTTEL RDKKQKVYSLFYRLDVVQ
 INSGSGGEYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDKKFNGTGPCQNVSTV
 QCTHGIKPVVSTQLLNGSLAEEVIIRSENITNNAKNILVQLNTSVQINCTRPNNNTVKSIRI
 GPGQAFYYTGDIIGDIRQAHCNVSKATWNETLGKVVQQLRKHFNGNTIIRFAQSSGGDLE
 VTTHSFNCGGEFFYCNTSGLFNSTW**PENG**T**MEGS**NGTITLPCRIKQIINMWQRIGQAMYAPP
 IQGVI RCVSNITGLILTRDGGSTNSTTETFRPGGDMRDNRSELYKYKVKIEPLGVAPTR
 CKRR

[00307] # Same as BG505_MD39_CP1.2_GRSF4_qLoops1, but adds three additional
 glycosylation sites (indicated by underline, bold). (This GRSF7 adds more glycans that
 BG505_MD39_GRSF7).

[00308] BG505_MD39_CP1.2_GRSF7_qLoops1

[00309] **GGN**SSSLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLK
 DTHWGIKQLQARVLAVEHYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWL
 QWDKEISNYTQIIYGLLEESQNQ**N**ESNEQDLGGNGSGGGSGSGNGSSGLWVTVYYGVPVWK

DAETTLFCASDAKAYETEKHNVWATHACVPTDPNSSEIHLENVTEEFNMWKNNMVEQMHEDI
ISLWDQSLKPCVKLTPLCVTLQCTNVTNNITD**NMT**GELKNCSFNMTTEL RDKKQKVYSLFYR
LDVVQ**INGSGGEYRL**INCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDKKFNGTGPCQ
NVSTVQCTHGIKPVVSTQLLLNGSLAEEVVIIRSENITNNAKNILVQLNT**SVQ**INCTRPNNN
TVKSIRIGPGQAFYYTGDIIGDIRQAHNCVSKATWNETLGKVVQQLRKHFNGNTIIRFAQSS
GGDLEVTTHSFNCGGEFFYCNTSGLFNSTW**PENGTMEGS**NGTITLPCRIKQIINMWQRIGQA
MYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGGGMRDNWRSELYKYKVVKIEPLG
VAPTRC**NRS**

[00310] # Cleavage-dependent BG505_MD39_GRSF4

[00311] but with "quiet" loops from BG505_MD39_CP1.2_GRSF4_qLoops1

[00312] BG505_MD39_GRSF4_qLoops1

[00313] AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNSSEIH
LENVTEEFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITD**NMT**GELK
NCSFNMTTEL RDKKQKVYSLFYRLDVVQ**INGSGGEYRL**INCNTSAITQACPKVSFEPIPIHY
CAPAGFAILKCKDKKFNGTGPCQNVSTVQCTHGIKPVVSTQ**LLLNGSLAEEVVIIRSENITN**
NAKNILVQLNT**SVQ**INCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHNCVSKATWNETL
GKVVQQLRKHFNGNTIIRFAQSSGGDLEVTTHSFNCGGEFFYCNTSGLFNSTW**PENGTMEGS**
NGTITLPCRIKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGG
GMRDNWRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVSLGFLGAAGSTMGA
ASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARVLAVEHYLRDQQLLG
IWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKN
NQSLLALDNGS

[00314] Type XI sequences

[00315] # BG505 SOSIP MD39 fused via GSGG linker to the Pyrococcus furiosus Ferritin
sequence from PDB: 2JD6. Compared to the ferritin sequence in 2JD6, the first position in
ferritin was mutated from M to G to contribute to the linker. Furthermore, another position
within ferritin was mutated from R to K, to eliminate a potential furin cleavage site.

[00316] BG505_SOSIP_MD39_2JD6

[00317] AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIH
LENVTEEFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELK
NCSFNMTTEL RDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVS
FEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQ**LLLNGSLAEEVVI**
IRSENITNNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHNCVS
KATWNETLGKVVQQLRKHFNGNTIIRFAQSSGGDLEVTTHSFNCGGEFFYCNTSGLFNSTWI
SNTSVQGSNSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGG
STNSTTETFRPGGGMRDNWRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVS
LGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARVL

AVEHYLRDQQLLGIWGC SGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
 GLEESQNQQEKNEQDLLALD**GS**GGLSERMLKALNDQLNRELYSAYLYFAMAAYFEDLGLEG
 FANWMKAQAEIIIIGHALRFYNYIYD**K**NGRVELDEIPKPPKEWESPLKAFAEAYEHEKFISKS
 IYELAALAEIIIKDYSTRAFLEWFINEQVEEEASVKKIILDKLKFAKDSPQILFMLDKELSARA
 PKLPGLLMQGGE**

[00318] # BG505 SOSIP MD39 fused via ASG linker to a variant of Dihydrolipoyl Transacetylase (E2p) from *Bacillus Stearothermophilus*. The E2p sequence was obtained from PDB: 1B5S, and the single unpaired cysteine in that sequence was either left in place, mutated to A, or mutated to T, as indicated by the "[C/A/T]" expression in the sequence below. Applicants have found that expression levels are superior when the unpaired cysteine is left intact or when it is mutated to T. The data in FIG. 20 correspond to the particle with an A at that position.

[00319] BG505 SOSIP MD39 E2p (also referred to as "MD39-1b5s")

[00320] AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQEIH
 LENVTEEFNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELK
 NCSFNMTTEL RDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVS
 FEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLNGSLAEIIIY
 IRSENITNNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNVS
 KATWNETLGKVVQQLRKHFGNNTIIRFAQSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWI
 SNTSVQGSNSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQGVIRCVSNTITGLILTRDGG
 STNSTTETFRPGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRAVGIGAVS
 LGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARVL
 AVEHYLRDQQLLGIWGC SGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
 GLEESQNQQEKNEQDLLALD**AS**GAAAKPATTEGEFFPETREKMSGIRRAIAKAMVHSHKHTAP
 HVTLMDEADVTKLVHRKKFKAIAAEKGIKLTFLPVVKALVSALREYPVLNT [C/A/T] ID
 DETEEIIQKHYYNIGIAADTDRGLLVPIKHADRKPIFALAQEINELA EKARDGKLTPEGEMK
 GASCTITNIGSAGGQWFTPVINHPEVAILGIGRIAEKPIVRDGEIVAAPMLALSLSFDHRMI
 DGATAQKALNHIKRLLSDPPELLLM**

[00321] Type XII Sequences

[00322] The first set in Type XII are cleavage-dependent, TM-anchored trimers, some with glycan masking.

[00323] #BG505 SOSIP MD39 as gp160 but without cytoplasmic domain

[00324] BG505_MD39_gp160_dCT

[00325] AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQEIH
 LENVTEEFNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELK
 NCSFNMTTEL RDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVS

FEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGLAESEVI
 IRSENITNNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNVS
 KATWNETLGKVVQQLRKHFNGNTIIRFAQSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWI
 SNTSVQGSNSTGSNDSITLPCRICKIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGG
 STNSTTETFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVS
 LGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARVL
 AVEHYLRDQQLLGIWGC SGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
 GLEESQNQQEKNEQDLLALDKWASLWNWFDISNWLWYIKIFIMIVGGLIGLRIVFAVLSVI
 HRVRQGYSPLS**

[00326] # This is BG505_MD39_gp160_dCT but adds 3 glycosylation sites (80, 241, 289). These 3 additional glycans are a subset of the 5 additional glycans present in BG505_MD39_GRSF4.

[00327] BG505_MD39_gp160_dCT_GRSF4.1

[00328] AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNSSEIH
 LENVTEEFNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELK
 NCSFNMTTEL RDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVS
 FEPIPIHYCAPAGFAILKCKDKKFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGLAESEVI
 IRSENITNNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNVS
 KATWNETLGKVVQQLRKHFNGNTIIRFAQSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWI
 SNTSVQGSNSTGSNDSITLPCRICKIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGG
 STNSTTETFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVS
 LGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARVL
 AVEHYLRDQQLLGIWGC SGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
 GLEESQNQQEKNEQDLLALDKWASLWNWFDISNWLWYIKIFIMIVGGLIGLRIVFAVLSVI
 HRVRQGYSPLS**

[00329] # This is BG505_MD39_gp160_dCT_GRSF4.1 plus the CT domain.

[00330] # Full-length gp160 including the Cytoplasmic domain.

[00331] BG505_MD39_gp160_GRSF4.1_m

[00332] AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNSSEIH
 LENVTEEFNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELK
 NCSFNMTTEL RDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVS
 FEPIPIHYCAPAGFAILKCKDKKFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGLAESEVI
 IRSENITNNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNVS
 KATWNETLGKVVQQLRKHFNGNTIIRFAQSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWI
 SNTSVQGSNSTGSNDSITLPCRICKIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGG
 STNSTTETFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVS
 LGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARVL

AVEHYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
 GLLEESQNQQEKNEQDLLALDKWASLWNWFDISNWLWYIKIFIMIVGGLIGLRIVFAVLSVI
 HRVRQGYSPLSFQTHTPNPRGLDRPERIEEEDGEQDRGRSTRLVSGFLALAWDDLRLSLCLFC
 YHRLRDFILIAARIVELLGHSSLKGLRLGWEGKYLWNLLAYWGRELKISAINLFDITIAIAV
 AEWTDREVIEIGQRLCRAFLHIPRRIRQGLERALL**

[00333] #Same as BG505_MD39_gp160_dCT but using linker and PDGFR TM domain instead of native TM domain, for better expression.

[00334] BG505_MD39_gp140-PDGFR

[00335] AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHL
 LENVTEEFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELK
 NCSFNMTTELDRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVS
 FEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLNGLSLAEEEVI
 IRSENITNNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNVS
 KATWNETLGKVVQQLRKHFNGNTIIRFAQSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWI
 SNTSVQGSNSTGSNDSITLPCRICKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGG
 STNSTTETFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVS
 LGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARVL
 AVEHYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
 GLLEESQNQQEKNEQDLLALDGGSGGGSGGSEQKLISEEDLGGSGGGSGGSSNAVGQDTQEIV
 VPHSLPFGVVVISAILALVVLTIISLIILIMLWQKKPR**

[00336] #Same as BG505_MD39_gp160_dCT_GRSF4.1 but using linker and PDGFR TM domain instead of native TM domain, for better expression. Same as BG505_MD39_gp140-PDGFR but adding 3 glycans at positions 80, 241, 289.

[00337] BG505_MD39_gp140-PDGFR_GRSF4.1_m

[00338] AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNSSEIHL
 LENVTEEFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELK
 NCSFNMTTELDRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVS
 FEPIPIHYCAPAGFAILKCKDKKFNGTGPCQNVSTVQCTHGIKPVVSTQLLNGLSLAEEEVI
 IRSENITNNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNVS
 KATWNETLGKVVQQLRKHFNGNTIIRFAQSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWI
 SNTSVQGSNSTGSNDSITLPCRICKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGG
 STNSTTETFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVS
 LGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARVL
 AVEHYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
 GLLEESQNQQEKNEQDLLALDGGSGGGSGGSEQKLISEEDLGGSGGGSGGSSNAVGQDTQEIV
 VPHSLPFGVVVISAILALVVLTIISLIILIMLWQKKPR**

[00339] The second set of Type XII sequences are cleavage-independent, TM-anchored trimers, some with glycan masking.

[00340] #BG505_MD39_link14 anchored to membrane by linker and PDGFR TM domain

[00341] BG505_MD39_gp140-PDGFR_link14

[00342] AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQEIH
LENVTEEFNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELK
NCSFNMTELRDCKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVS
FEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEEVEI
IRSENITNNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNVS
KATWNETLGKVVQQLRKHFNGNTIIRFAQSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWI
SNTSVQGSNSTGSNDSITLPCRICKQIINMWQRIGQAMYAPPIQGVIKRVSNITGLILTRDGG
STNSTTETFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGSHSGSGSGSGGHA
AVGIGAVSLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTWGI
KQLQARVLAVEHYLRDQQLLGIWGC SGK LICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEI
SNYTQIIYGLLEESQNQQEKNEQDLLALDGGSGSGSGGSEQKLISEEDLGGSGSGSGSNAVG
QDTQEVIVVPHSLPFKVVVISAILALVVLTIISLIILIMLWQKKPR**

[00343] #Cleavage-independent, link14 version of BG505_MD39_gp160_dCT.

[00344] BG505_MD39_gp160-dCT_link14

[00345] AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQEIH
LENVTEEFNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELK
NCSFNMTELRDCKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVS
FEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEEVEI
IRSENITNNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNVS
KATWNETLGKVVQQLRKHFNGNTIIRFAQSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWI
SNTSVQGSNSTGSNDSITLPCRICKQIINMWQRIGQAMYAPPIQGVIKRVSNITGLILTRDGG
STNSTTETFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGSHSGSGSGSGGHA
AVGIGAVSLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTWGI
KQLQARVLAVEHYLRDQQLLGIWGC SGK LICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEI
SNYTQIIYGLLEESQNQQEKNEQDLLALDKWASLWNWFDISNWLWYIKIFIMIVGGLIGLRI
VFAVLSVIHRVRQGYSPS**

[00346] #Same as BG505_MD39_gp160-dCT_link14 but adding 3 glycans, at positions 80, 241, 289.

[00347] BG505_MD39_gp160-dCT_link14_GRSF4.1

[00348] AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPN**SS**EIH
LENVTEEFNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELK
NCSFNMTELRDCKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVS
FEPIPIHYCAPAGFAILKCKDKKFNGTGPC**Q**NVSTVQCTHGIKPVVSTQLLLNGSLAEEVEI

IRSENITNNAKNILVQLNTSVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNVS
 KATWNETLGKVVQQLRKHFNGNTIIRFAQSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWI
 SNTSVQGSNSTGSNDSITLPCRICKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGG
 STNSTTETFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGSHSGSGSGSGGHA
 AVGIGAVSLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDPHWDGI
 KQLQARVLAVEHYLRDQQLLGIWGC SGKLCCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEI
 SNYTQIIYGLLEESQNQQEKNEQDLLALDKWASLWNWFDISNWLWYIKIFIMIVGGLIGLRI
 VFAVL SVIHRVRQGYSPLS**

[00349] #Same as BG505_MD39_gp160-dCT_link14_GRSF4.1 but using linker and PDGFR TM domain instead of native TM domain, for better expression.

[00350] BG505_MD39_gp140-PDGFR_link14_GRSF4.1

[00351] AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNSSEIH
 LENVTEEFNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELK
 NCSFNMTTELDRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVS
 FEPIPIHYCAPAGFAILKCKDKKFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEVEI
 IIRSENITNNAKNILVQLNTSVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNVS
 KATWNETLGKVVQQLRKHFNGNTIIRFAQSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWI
 SNTSVQGSNSTGSNDSITLPCRICKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGG
 STNSTTETFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGSHSGSGSGSGGHA
 AVGIGAVSLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDPHWDGI
 KQLQARVLAVEHYLRDQQLLGIWGC SGKLCCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEI
 SNYTQIIYGLLEESQNQQEKNEQDLLALDGGSGSGSGGSEQKLISEEDLGGSGSGSGSNAVG
 QDTQEVIVVPHSLPFKVVVISAILALVVLTIISLIILIMLWQKKPR**

[00352] # BG505_MD39_CP1.2_GRSF4 plus linker and PDGFR TM domain.

[00353] # Similar to BG505_MD39_gp140-PDGFR_link14_GRSF4.1 but is CP1.2 instead of link14 and GRSF4.0 instead of GRSF4.1

[00354] BG505_MD39_gp140-PDGFR_CP1.2_GRSF4.0

[00355] VSLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDPHWDGI
 IKQLQARVLAVEHYLRDQQLLGIWGC SGKLCCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEI
 ISNYTQIIYGLLEESQNQQEKNEQDLGGNGSGGGSGSGGNGSSGLWVTVYYGVPVWKDAETT
 LFCASDAKAYETEKHNWATHACVPTDPNSSEIHLENVTEEFNMWKNMVEQMHEDIISLWD
 QSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELKNCSFNMTTELDRDKKQKVYSLFYRLDVVQ
 INENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDKKFNGTG
 PCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEVEIIRSENITNNAKNILVQLNTSVQINCTRP
 NNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNVSKATWNETLGKVVQQLRKHFNGNTIIRFA
 QSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQGSNSTGSNDSITLPCRICKQIIN
 MWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGGGDMRDNRSELYKYK

VVKIEPLGVAPTRCKRRGGGSGGSGGSEQKLISEEDLGGSGGSGGSNAVGQDTQEVIIVPHS
LPFKVVVISAILALVVLTIISLIILIMLWQKKPR**

[00356] # Different Circular permutation than CP1.2; this one uses native gp41 C-term so might be better for membrane presentation. Indeed, CP2 doesn't need flexible linker to the membrane. This version is gp160 minus the cytoplasmic domain and has additional glycosylation sites added to positions 80, 241, 289.

[00357] BG505_MD39_gp160-dCT_CP2_GRSF4.1

[00358] VSLGFLGAAGSTMGAAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWG
IKQLQARVLAVEHYLRDQQLLGIWGCSGKLICCTGGSGVTVYYGVPVWKDAETTLFCASDAK
AYETEKHNVWATHACVPTDPNSSEIHLENVTEEFNMWKNNMVEQMHEDIISLWDQSLKPCVK
LTPLCVTLQCTNVNTNITDDMRGELKNCSFNMTTEL RDKKQKVYSLFYRLDVVQINENQGNR
SNNSNKEYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDKKFNGTGPCQNVSTV
QCTHGIKPVVSTQLLLNGSLAEEEVIIRSENITNNAKNILVQLNTSVQINCTRPNNTVKS
RIGPGQAFYYTGDIIGDIRQAHCNVSKATWNETLGKVVKQLRKHFNGNTIIRFAQSSGGDLE
VTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQGSNSTGSNDSITLPCRIKQIINMWQRIGQA
MYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGGDMRDNRSELYKYKVVKIEPLG
VAPTRCKRGGSGSGVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEK
NEQDLLALDKWASLWNWFDISNWLWYIKIFIMIVGGLIGLRIVFAVLSVIHRVRQGYSPLS*
*

[00359] # BG505_MD39_gp160-dCT_CP2_GRSF4.1 but with two extra glycans (NQSLALDNGS) and with linker and PDGFR TM instead of native TM.

[00360] BG505_MD39_gp140-PDGFR_CP2_GRSF4.2

[00361] VSLGFLGAAGSTMGAAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWG
IKQLQARVLAVEHYLRDQQLLGIWGCSGKLICCTGGSGVTVYYGVPVWKDAETTLFCASDAK
AYETEKHNVWATHACVPTDPNSSEIHLENVTEEFNMWKNNMVEQMHEDIISLWDQSLKPCVK
LTPLCVTLQCTNVNTNITDDMRGELKNCSFNMTTEL RDKKQKVYSLFYRLDVVQINENQGNR
SNNSNKEYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDKKFNGTGPCQNVSTV
QCTHGIKPVVSTQLLLNGSLAEEEVIIRSENITNNAKNILVQLNTSVQINCTRPNNTVKS
RIGPGQAFYYTGDIIGDIRQAHCNVSKATWNETLGKVVKQLRKHFNGNTIIRFAQSSGGDLE
VTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQGSNSTGSNDSITLPCRIKQIINMWQRIGQA
MYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGGDMRDNRSELYKYKVVKIEPLG
VAPTRCKRGGSGSGVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEK
NNQSLALDNGSGGGSGGSGGSEQKLISEEDLGGSGGSGGSNAVGQDTQEVIIVPHSLPFKV
VVISAILALVVLTIISLIILIMLWQKKPR**

[00362] # BG505_MD39_gp160-dCT_CP2_GRSF4.1 but with CT included, so this is full-length gp160.

[00363] BG505_MD39_gp160_CP2_GRSF4.1

[00364] VSLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWG
IKQLQARVLAVEHYLRDQQLLGIWGCSGKLICTGGSGVTVYYGVPVWKDAETTLFCASDAK
AYETEKHNVWATHACVPTDPNSSEIHLENVTEEFNMWKNNMVEQMHEDIISLWDQSLKPCVK
LTPLCVTLQCTNVTNNITDDMRGELKNCSFNMTTEL RDKKQKVYSLFYRLDVQINENQGNR
SNNSNKEYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDKKFNGTGPCQNVSTV
QCTHGIKPVVSTQLLLNGSLAEEEVIIIRSENITNNAKNILVQLNTSVQINCTRPNNTVKSI
RIGPGQAFYYTGDIIGDIRQAHCNVSKATWNETLGKVVKQLRKHFGNNTIIRFAQSSGGDLE
VTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQGSNSTGSNDSITLPCRIKQIINMWQRIGQA
MYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGGDMRDNWRSELYKYKVVKIEPLG
VAPTRCKRGGSGSGVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQOEK
NEQDLLALDKWASLWNWFDISNWLWYIKIFIMIVGGLIGLRIVFAVLSVIHRVRQGYSPLSF
QHTPNPRGLDRPERIEEEDGEQDRGRSTRLVSGFLALAWDDLRLCLFCYHRLRDFILIAA
RIVELLGHSSSLKGLRLGWEGLKYLWNLAYWGRELKISAINLFDITIAIAVAEWTDREVIEIGQ
RLCRAFLHIPRRIRQGLERALL**

[00365] In one embodiment, the nucleic acids of the present invention may be delivered as a therapeutic mRNA.

[00366] Provided herein are isolated nucleic acids (e.g., modified mRNAs encoding a peptide described herein) comprising a translatable region and at least two different nucleoside modifications, wherein the nucleic acid exhibits reduced degradation in a cell into which the nucleic acid is introduced, relative to a corresponding unmodified nucleic acid. For example, the degradation rate of the nucleic acid is reduced by at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90%, compared to the degradation rate of the corresponding unmodified nucleic acid. In certain embodiments, the nucleic acid comprises RNA, DNA, TNA, GNA, or a hybrid thereof. In certain embodiments, the nucleic acid comprises messenger RNA (mRNA). In certain embodiments, the mRNA does not substantially induce an innate immune response of the cell into which the mRNA is introduced. In certain embodiments, the mRNA comprises at least one nucleoside selected from the group consisting of pyridin-4-one ribonucleoside, 5-aza-uridine, 2-thio-5-aza-uridine, 2-thiouridine, 4-thio-pseudouridine, 2-thio-pseudouridine, 5-hydroxyuridine, 3-methyluridine, 5-carboxymethyl-uridine, 1-carboxymethyl-pseudouridine, 5-propynyl-uridine, 1-propynyl-pseudouridine, 5-taurinomethyluridine, 1-taurinomethyl-pseudouridine, 5-taurinomethyl-2-thio-uridine, 1-taurinomethyl-4-thio-uridine, 5-methyl-uridine, 1-methyl-pseudouridine, 4-thio-1-methyl-pseudouridine, 2-thio-1-methyl-pseudouridine, 1-methyl-1-deaza-pseudouridine, 2-thio-1-methyl-1-deaza-pseudouridine, dihydrouridine,

dihydropseudouridine, 2-thio-dihydrouridine, 2-thio-dihydropseudouridine, 2-methoxyuridine, 2-methoxy-4-thio-uridine, 4-methoxy-pseudouridine, and 4-methoxy-2-thio-pseudouridine. In certain embodiments, the mRNA comprises at least one nucleoside selected from the group consisting of 5-aza-cytidine, pseudoisocytidine, 3-methyl-cytidine, N4-acetylcytidine, 5-formylcytidine, N4-methylcytidine, 5-hydroxymethylcytidine, 1-methyl-pseudoisocytidine, pyrrolo-cytidine, pyrrolo-pseudoisocytidine, 2-thio-cytidine, 2-thio-5-methyl-cytidine, 4-thio-pseudoisocytidine, 4-thio-1-methyl-pseudoisocytidine, 4-thio-1-methyl-1-deaza-pseudoisocytidine, 1-methyl-1-deaza-pseudoisocytidine, zebularine, 5-aza-zebularine, 5-methyl-zebularine, 5-aza-2-thio-zebularine, 2-thio-zebularine, 2-methoxy-cytidine, 2-methoxy-5-methyl-cytidine, 4-methoxy-pseudoisocytidine, and 4-methoxy-1-methyl-pseudoisocytidine. In other embodiments, the mRNA comprises at least one nucleoside selected from the group consisting of 2-aminopurine, 2,6-diaminopurine, 7-deaza-adenine, 7-deaza-8-aza-adenine, 7-deaza-2-aminopurine, 7-deaza-8-aza-2-aminopurine, 7-deaza-2,6-diaminopurine, 7-deaza-8-aza-2,6-diaminopurine, 1-methyladenosine, N6-methyladenosine, N6-isopentenyladenosine, N6-(cis-hydroxyisopentenyl)adenosine, 2-methylthio-N6-(cis-hydroxyisopentenyl) adenosine, N6-glycinylicarbamoyladenosine, N6-threonylicarbamoyladenosine, 2-methylthio-N6-threonyl carbamoyladenosine, N6,N6-dimethyladenosine, 7-methyladenine, 2-methylthio-adenine, and 2-methoxy-adenine. In yet other embodiments, the mRNA comprises at least one nucleoside selected from the group consisting of inosine, 1-methyl-inosine, wyosine, wybutosine, 7-deaza-guanosine, 7-deaza-8-aza-guanosine, 6-thio-guanosine, 6-thio-7-deaza-guanosine, 6-thio-7-deaza-8-aza-guanosine, 7-methyl-guanosine, 6-thio-7-methyl-guanosine, 7-methylinosine, 6-methoxy-guanosine, 1-methylguanosine, N2-methylguanosine, N2,N2-dimethylguanosine, 8-oxo-guanosine, 7-methyl-8-oxo-guanosine, 1-methyl-6-thio-guanosine, N2-methyl-6-thio-guanosine, and N2,N2-dimethyl-6-thio-guanosine.

[00367] In some embodiments, the nucleic acids provided herein comprise a 5' untranslated region (UTR) and/or a 3'UTR, wherein each of the two different nucleoside modifications are independently present in the 5'UTR and/or 3'UTR. In some embodiments, nucleic acids are provided herein, wherein at least one of the two different nucleoside modifications are present in the translatable region. In some embodiments, nucleic acids provided herein are capable of binding to at least one polypeptide that prevents or reduces an innate immune response of a cell into which the nucleic acid is introduced.

[00368] Further provided herein are isolated nucleic acids (e.g., modified mRNAs described herein) comprising (i) a translatable region encoding a peptide described herein, (ii) at least one nucleoside modification, and (iii) at least one intronic nucleotide sequence capable of being excised from the nucleic acid.

[00369] Further provided herein are isolated nucleic acids (e.g., modified mRNAs described herein) comprising (i) a translatable region encoding a peptide described herein, (ii) at least two different nucleoside modifications, and (iii) a degradation domain.

[00370] Further provided herein are non-enzymatically synthesized nucleic acids (e.g., modified mRNAs described herein) comprising at least one nucleoside modification, and comprising a translatable region encoding a peptide described herein. In certain embodiments, the non-enzymatically synthesized mRNA comprises at least two different nucleoside modifications.

[00371] Further provided herein are isolated nucleic acids (e.g., modified mRNAs described herein) comprising a noncoding region and at least one nucleoside modification that reduces an innate immune response of a cell into which the nucleic acid is introduced, wherein the nucleic acid sequesters one or more translational machinery components. In certain embodiments, the isolated nucleic acids comprising a noncoding region and at least one nucleoside modification described herein are provided in an amount effective to reduce protein expression in the cell. In certain embodiments, the translational machinery component is a ribosomal protein or a transfer RNA (tRNA). In certain embodiments, the nucleic acid comprises a small nucleolar RNA (sno-RNA), microRNA (miRNA), small interfering RNA (siRNA) or Piwi-interacting RNA (piRNA).

[00372] Further provided herein are isolated nucleic acids (e.g., modified mRNAs described herein) comprising (i) a first translatable region, (ii) at least one nucleoside modification, and (iii) an internal ribosome entry site (IRES). In certain embodiments, the IRES is obtained from a picornavirus, a pest virus, a polio virus, an encephalomyocarditis virus, a foot-and-mouth disease virus, a hepatitis C virus, a classical swine fever virus, a murine leukemia virus, a simian immune deficiency virus or a cricket paralysis virus. In certain embodiments, the isolated nucleic acid further comprises a second translatable region. In certain embodiments, the isolated nucleic acid further comprises a Kozak sequence. In some embodiments, the first translatable region encodes a peptide described herein. In some embodiments, the second translatable region encodes peptide described herein. In some embodiments, the first and the second translatable regions encode peptides described herein.

[00373] Provided herein are pharmaceutical compositions comprising: (i) an effective amount of a synthetic messenger ribonucleic acid (mRNA) encoding peptide described herein; and (ii) a pharmaceutically acceptable carrier, wherein i) the mRNA comprises pseudouridine, 5'methyl-cytidine, or a combination thereof, or ii) the mRNA does not comprise a substantial amount of a nucleotide or nucleotides selected from the group consisting of uridine, cytidine, and a combination of uridine and cytidine, and wherein the composition is suitable for repeated administration (e.g., intravenous administration) to a mammalian subject in need thereof. In some embodiments,

[00374] Further provided herein are pharmaceutical compositions comprising and/or consisting essentially of: (i) an effective amount of a synthetic messenger ribonucleic acid (mRNA) encoding peptide described herein; (ii) a cell penetration agent; and (iii) a pharmaceutically acceptable carrier, wherein i) the mRNA comprises pseudouridine, 5'methyl-cytidine or a combination thereof, or ii) the mRNA does not comprise a substantial amount of a nucleotide or nucleotides selected from the group consisting of uridine, cytidine, and a combination of uridine and cytidine, and wherein the composition is suitable for repeated administration (e.g., intravenous administration) to an animal (e.g., mammalian) subject in need thereof.

[00375] This invention provides nucleic acids, including RNAs such as mRNAs that contain one or more modified nucleosides (termed "modified nucleic acids"), which have useful properties including the lack of a substantial induction of the innate immune response of a cell into which the mRNA is introduced. Because these modified nucleic acids enhance the efficiency of protein production, intracellular retention of nucleic acids, and viability of contacted cells, as well as possess reduced immunogenicity, these nucleic acids having these properties are termed "enhanced nucleic acids" herein.

[00376] The term "nucleic acid," in its broadest sense, includes any compound and/or substance that is or can be incorporated into an oligonucleotide chain. Exemplary nucleic acids for use in accordance with the present invention include, but are not limited to, one or more of DNA, RNA, hybrids thereof, RNAi-inducing agents, RNAi agents, siRNAs, shRNAs, miRNAs, antisense RNAs, ribozymes, catalytic DNA, RNAs that induce triple helix formation, aptamers, vectors, etc., described in detail herein.

[00377] Provided are modified nucleic acids containing a translatable region encoding a peptide described herein, and one, two, or more than two different nucleoside modifications. In some embodiments, the modified nucleic acid exhibits reduced degradation in a cell into

which the nucleic acid is introduced, relative to a corresponding unmodified nucleic acid. For example, the degradation rate of the nucleic acid is reduced by at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90%, compared to the degradation rate of the corresponding unmodified nucleic acid. Exemplary nucleic acids include ribonucleic acids (RNAs), deoxyribonucleic acids (DNAs), threose nucleic acids (TNAs), glycol nucleic acids (GNAs), peptide nucleic acids (PNAs), locked nucleic acids (LNAs) or a hybrid thereof. In preferred embodiments, the modified nucleic acid includes messenger RNAs (mRNAs). As described herein, the nucleic acids of the invention do not substantially induce an innate immune response of a cell into which the mRNA is introduced.

[00378] In some embodiments, modified nucleosides include pyridin-4-one ribonucleoside, 5-aza-uridine, 2-thio-5-aza-uridine, 2-thiouridine, 4-thio-pseudouridine, 2-thio-pseudouridine, 5-hydroxyuridine, 3-methyluridine, 5-carboxymethyl-uridine, 1-carboxymethyl-pseudouridine, 5-propynyl-uridine, 1-propynyl-pseudouridine, 5-taurinomethyluridine, 1-taurinomethyl-pseudouridine, 5-taurinomethyl-2-thio-uridine, 1-taurinomethyl-4-thio-uridine, 5-methyl-uridine, 1-methyl-pseudouridine, 4-thio-1-methyl-pseudouridine, 2-thio-1-methyl-pseudouridine, 1-methyl-1-deaza-pseudouridine, 2-thio-1-methyl-1-deaza-pseudouridine, dihydrouridine, dihydropseudouridine, 2-thio-dihydrouridine, 2-thio-dihydropseudouridine, 2-methoxyuridine, 2-methoxy-4-thio-uridine, 4-methoxy-pseudouridine, and 4-methoxy-2-thio-pseudouridine.

[00379] In some embodiments, modified nucleosides include 5-aza-cytidine, pseudoisocytidine, 3-methyl-cytidine, N4-acetylcytidine, 5-formylcytidine, N4-methylcytidine, 5-hydroxymethylcytidine, 1-methyl-pseudoisocytidine, pyrrolo-cytidine, pyrrolo-pseudoisocytidine, 2-thio-cytidine, 2-thio-5-methyl-cytidine, 4-thio-pseudoisocytidine, 4-thio-1-methyl-pseudoisocytidine, 4-thio-1-methyl-1-deaza-pseudoisocytidine, 1-methyl-1-deaza-pseudoisocytidine, zebularine, 5-aza-zebularine, 5-methyl-zebularine, 5-aza-2-thio-zebularine, 2-thio-zebularine, 2-methoxy-cytidine, 2-methoxy-5-methyl-cytidine, 4-methoxy-pseudoisocytidine, and 4-methoxy-1-methyl-pseudoisocytidine.

[00380] In other embodiments, modified nucleosides include 2-aminopurine, 2,6-diaminopurine, 7-deaza-adenine, 7-deaza-8-aza-adenine, 7-deaza-2-aminopurine, 7-deaza-8-aza-2-aminopurine, 7-deaza-2,6-diaminopurine, 7-deaza-8-aza-2,6-diaminopurine, 1-methyladenosine, N6-methyladenosine, N6-isopentenyladenosine, N6-(cis-hydroxyisopentenyl)adenosine, 2-methylthio-N6-(cis-hydroxyisopentenyl) adenosine, N6-

glycylcarbamoyladenine, N6-threonylcarbamoyladenine, 2-methylthio-N6-threonylcarbamoyladenine, N6,N6-dimethyladenine, 7-methyladenine, 2-methylthio-adenine, and 2-methoxy-adenine.

[00381] In certain embodiments it is desirable to intracellularly degrade a modified nucleic acid introduced into the cell, for example if precise timing of protein production is desired. Thus, the invention provides a modified nucleic acid containing a degradation domain, which is capable of being acted on in a directed manner within a cell.

[00382] In other embodiments, modified nucleosides include inosine, 1-methyl-inosine, wyosine, wybutosine, 7-deaza-guanosine, 7-deaza-8-aza-guanosine, 6-thio-guanosine, 6-thio-7-deaza-guanosine, 6-thio-7-deaza-8-aza-guanosine, 7-methyl-guanosine, 6-thio-7-methyl-guanosine, 7-methylinosine, 6-methoxy-guanosine, 1-methylguanosine, N2-methylguanosine, N2,N2-dimethylguanosine, 8-oxo-guanosine, 7-methyl-8-oxo-guanosine, 1-methyl-6-thio-guanosine, N2-methyl-6-thio-guanosine, and N2,N2-dimethyl-6-thio-guanosine.

[00383] Other components of nucleic acid are optional, and are beneficial in some embodiments. For example, a 5' untranslated region (UTR) and/or a 3'UTR are provided, wherein either or both may independently contain one or more different nucleoside modifications. In such embodiments, nucleoside modifications may also be present in the translatable region. Also provided are nucleic acids containing a Kozak sequence.

[00384] Further, nucleic acids encoding a peptide described herein, and containing an internal ribosome entry site (IRES) are provided herein. An IRES may act as the sole ribosome binding site, or may serve as one of multiple ribosome binding sites of an mRNA. An mRNA containing more than one functional ribosome binding site may encode several peptides or polypeptides that are translated independently by the ribosomes ("multicistronic mRNA"). When nucleic acids are provided with an IRES, further optionally provided is a second translatable region. Examples of IRES sequences that can be used according to the invention include without limitation, those from picornaviruses (e.g., FMDV), pest viruses (CFFV), polio viruses (PV), encephalomyocarditis viruses (ECMV), foot-and-mouth disease viruses (FMDV), hepatitis C viruses (HCV), classical swine fever viruses (CSFV), murine leukemia virus (MLV), simian immune deficiency viruses (SIV) or cricket paralysis viruses (CrPV).

[00385] The therapeutic mRNAs as described, for example, in US Patent Nos. 9,464,124; 9,447,164; 9,428,535; 9,334,328; 9,303,079; 9,301,993; 9,295,689; 9,283,287; 9,271,996; 9,255,129; 9,254,311; 9,233,141; 9,221,891; 9,220,792; 9,220,755; 9,216,205; 9,192,651;

9,186,372; 9,181,319; 9,149,506; 9,114,113; 9,107,886; 9,095,552; 9,089,604; 9,061,059; 9,050,297; 8,999,380; 8,980,864; 8,822,663; 8,754,062; 8,710,200; 8,680,069 and 8,664,194 may be utilized for the present invention.

[00386] Methods for the chemical conjugation of polypeptides, carbohydrates, and/or lipids are well known in the art (see, for example, Hermanson. *Bioconjugate Techniques* (Academic Press; 1992); Aslam and Dent, eds. *Bioconjugation: Protein coupling Techniques for the Biomedical Sciences* (MacMillan: 1998); and Wong *Chemistry of Protein Conjugation and Cross-linking* (CRC Press: 1991)). For instance, primary amino groups may be incorporated by reaction with ethylenediamine in the presence of sodium cyanoborohydride and sulfhydryls may be introduced by reaction of cysteamine dihydrochloride followed by reduction with a standard disulfide reducing agent. Heterobifunctional crosslinkers, such as, for example, sulfosuccinimidyl (4-iodoacetyl) aminobenzoate, which link the epsilon amino group on the D-lysine residues of copolymers of D-lysine and D-glutamate to a sulfhydryl side chain from an amino terminal cysteine residue on the peptide to be coupled, may be used as well. Chemical conjugation also includes anything covalently bonded directly via side chain bonds or via a linker or spacer group.

[00387] The nanoparticle formulations may be a carbohydrate nanoparticle comprising a carbohydrate carrier and a modified nucleic acid molecule (e.g., mmRNA). As a non-limiting example, the carbohydrate carrier may include, but is not limited to, an anhydride-modified phytoglycogen or glycogen-type material, phytoglycogen octenyl succinate, phytoglycogen beta-dextrin, anhydride-modified phytoglycogen beta-dextrin. (See e.g., International Publication No. WO2012109121; herein incorporated by reference in its entirety).

[00388] Lipid nanoparticle formulations may be improved by replacing the cationic lipid with a biodegradable cationic lipid which is known as a rapidly eliminated lipid nanoparticle (reLNP). Ionizable cationic lipids, such as, but not limited to, DLinDMA, DLin-KC2-DMA, and DLin-MC3-DMA, have been shown to accumulate in plasma and tissues over time and may be a potential source of toxicity. The rapid metabolism of the rapidly eliminated lipids can improve the tolerability and therapeutic index of the lipid nanoparticles by an order of magnitude from a 1 mg/kg dose to a 10 mg/kg dose in rat. Inclusion of an enzymatically degraded ester linkage can improve the degradation and metabolism profile of the cationic component, while still maintaining the activity of the reLNP formulation. The ester linkage

can be internally located within the lipid chain or it may be terminally located at the terminal end of the lipid chain. The internal ester linkage may replace any carbon in the lipid chain.

[00389] The average diameter of the nanoparticle employed in the compositions of the invention can be at least one member selected from the group consisting of about 20 nanometers, about 25 nanometers, about 30 nanometers, about 40 nanometers, about 50 nanometers, about 75 nanometers, about 100 nanometers, about 125 nanometers, about 150 nanometers, about 175 nanometers and about 200 nanometers. In another embodiment, the average diameter of the particle is at least one member selected from the group consisting of between about 10 to about 200 nanometers, between about 0.5 to about 5 microns and between about 5 to about 10 microns. In another embodiment, the average diameter of the microparticle is selected from the group consisting of about 0.1 μm , about 0.2 μm , about 0.4 μm , about 0.5 μm , about 1 μm and about 2 μm .

[00390] Nanoparticles for use in the compositions of the invention can be made from lipids or other fatty acids (see, for example, U.S. Pat. Nos. 5,709,879; 6,342,226; 6,090,406; Lian, et al., J. of Pharma. Sci. 90:667-680 (2001) and van Slooten, et al., Pharm Res. 17:42-48 (2000)) and non-lipid compositions (see, for example, Kreuter, J. Anat. 189:503-505 (1996), the teachings of all of which are hereby incorporated by reference in their entirety). The compositions can be bilayer or multilamellar liposomes and phospholipid based. Polymerized nanoparticles, as described, for example, in U.S. Pat. No. 7,285,289, the teachings of which are incorporated by reference in their entirety.

[00391] Metallic oxide nanoparticles for use in the compositions of the invention can be chemically substituted with at least one reactive moiety capable of forming a thioether bond employing conventionally techniques as described herein and in U.S. Pat. No. 6,086,881, the teachings of which are hereby incorporated by reference in their entirety. The antigen described herein can be coupled in a single step onto the metallic oxide particles by the formation of at least one thioether bond or it may be synthesized or assembled stepwise onto the metallic oxide particles after the initial thioether bond formation. The chemical derivatization reagents for the metallic oxide particles can include organosilane reagents that provide thioalkane functionality or other groups that may readily be converted into thiols or thiol-reactive moieties. Organosilane reagents which may be utilized for this purpose may be, but are not limited to, 3-mercaptopropyltrimethoxysilane, 3-aminopropyltriethoxysilane, 3-iodopropyltrimethoxysilane, 2-chloroethyltrichlorosilane, 3-glycidoxypropyltrimethoxysilane, vinyltrichlorosilane and 3-acryloxypropyltrimethoxysilane.

Moieties that include one or more disulfide components may also be joined to the metallic oxide particle surface and thereby provide the corresponding reactive moiety able to enter into and form a thioether bond and juncture. Exemplary nanoparticles for use in the compositions of the invention include at least one member selected from the group consisting of poly (d,l-lactide-co-glycolide, also referred to as "poly(lactic-co-glycolic acid) and bisacyloxypropylcysteine.

[00392] Nanoparticles for use in the compositions of the invention can be made of inorganic material. Nanoparticles for use in the compositions of the invention can be made of a polymer material, such as at least one member selected from the group consisting of polystyrene, brominated polystyrene, polyacrylic acid, polyacrylonitrile, polyamide, polyacrylamide, polyacrolein, polybutadiene, polycaprolactone, polycarbonate, polyester, polyethylene, polyethylene terephthalate, polydimethylsiloxane, polyisoprene, polyurethane, polyvinylacetate, polyvinylchloride, polyvinylpyridine, polyvinylbenzylchloride, polyvinyltoluene, polyvinylidene chloride, polydivinylbenzene, polymethylmethacrylate, polylactide, polyglycolide, poly(lactide-co-glycolide), polyanhydride, polyorthoester, polyphosphazene, polyphosphazene, a carbohydrate, carboxymethyl cellulose, hydroxyethyl cellulose, agar, gel, proteinaceous polymer, polypeptide, eukaryotic and prokaryotic cells, viruses, lipid, metal, resin, latex, rubber, silicone (e.g., polydimethyldiphenyl siloxane), glass, ceramic, charcoal, kaolinite and bentonite.

[00393] It is noted that these therapeutics may be a chemical compound, a composition which may comprise a polypeptide of the present invention and/or antibody elicited by such a chemical compound and/or portion thereof or a pharmaceutically acceptable salt or a composition which may comprise a polypeptide of the invention, and may be administered alone or as an active ingredient in combination with pharmaceutically acceptable carriers, diluents, and vehicles, as well as other active ingredients.

[00394] The compounds or compositions may be administered orally, subcutaneously or parenterally including intravenous, intraarterial, intramuscular, intraperitoneally, and intranasal administration as well as intrathecal and infusion techniques.

[00395] It is noted that humans are treated generally longer than the mice or other experimental animals which treatment has a length proportional to the length of the disease process and drug effectiveness. The doses may be single doses or multiple doses over a period of several days, but single doses are preferred. Thus, one may scale up from animal

experiments, e.g., rats, mice, and the like, to humans, by techniques from this disclosure and documents cited herein and the knowledge in the art, without undue experimentation.

[00396] The treatment generally has a length proportional to the length of the disease process and drug effectiveness and the patient being treated.

[00397] When administering a therapeutic of the present invention parenterally, it will generally be formulated in a unit dosage injectable form (solution, suspension, emulsion). The pharmaceutical formulations suitable for injection include sterile aqueous solutions or dispersions and sterile powders for reconstitution into sterile injectable solutions or dispersions. The carrier may be a solvent or dispersing medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, liquid polyethylene glycol, and the like), suitable mixtures thereof, and vegetable oils.

[00398] Proper fluidity may be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Nonaqueous vehicles such as cottonseed oil, sesame oil, olive oil, soybean oil, corn oil, sunflower oil, or peanut oil and esters, such as isopropyl myristate, may also be used as solvent systems for compound compositions.

[00399] Additionally, various additives which enhance the stability, sterility, and isotonicity of the compositions, including antimicrobial preservatives, antioxidants, chelating agents, and buffers, may be added. Prevention of the action of microorganisms may be ensured by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, sorbic acid, and the like. In many cases, it will be desirable to include isotonic agents, for example, sugars, sodium chloride, and the like. Prolonged absorption of the injectable pharmaceutical form may be brought about by the use of agents delaying absorption, for example, aluminum monostearate and gelatin. According to the present invention, however, any vehicle, diluent, or additive used would have to be compatible with the compounds.

[00400] Sterile injectable solutions may be prepared by incorporating the compounds utilized in practicing the present invention in the required amount of the appropriate solvent with various amounts of the other ingredients, as desired.

[00401] A pharmacological formulation of the present invention, e.g., which may comprise a therapeutic compound or polypeptide of the present invention, may be administered to the patient in an injectable formulation containing any compatible carrier, such as various vehicles, adjuvants, additives, and diluents; or the compounds utilized in the present invention may be administered parenterally to the patient in the form of slow-release

subcutaneous implants or targeted delivery systems such as monoclonal antibodies, iontophoretic, polymer matrices, liposomes, and microspheres.

[00402] A pharmacological formulation of the compound and composition which may comprise a polypeptide utilized in the present invention may be administered orally to the patient. Conventional methods such as administering the compounds in tablets, suspensions, solutions, emulsions, capsules, powders, syrups and the like are usable. Known techniques, which deliver the compound orally or intravenously and retain the biological activity, are preferred.

[00403] In one embodiment, a formulation of the present invention may be administered initially, and thereafter maintained by further administration. For instance, a formulation of the invention may be administered in one type of composition and thereafter further administered in a different or the same type of composition. For example, a formulation of the invention may be administered by intravenous injection to bring blood levels to a suitable level. The patient's levels are then maintained by an oral dosage form, although other forms of administration, dependent upon the patient's condition, may be used. In the instance of a vaccine composition, the vaccine may be administered as a single dose, or the vaccine may incorporate set booster doses. For example, booster doses may comprise variants in order to provide protection against multiple clades of HIV.

[00404] The quantity to be administered will vary for the patient being treated and whether the administration is for treatment or prevention and will vary from a few micrograms to a few milligrams for an average 70 kg patient, e.g., 5 micrograms to 5 milligrams such as 500 micrograms, or about 100 ng/kg of body weight to 100 mg/kg of body weight per administration and preferably will be from 10 pg/kg to 10 mg/kg per administration. Typically, however, the antigen is present in an amount on, the order of micrograms to milligrams, or, about 0.001 to about 20 wt %, preferably about 0.01 to about 10 wt %, and most preferably about 0.05 to about 5 wt %.

[00405] Of course, for any composition to be administered to an animal or human, including the components thereof, and for any particular method of administration, it is preferred to determine therefor: toxicity, such as by determining the lethal dose (LD) and LD₅₀ in a suitable animal model e.g., rodent such as mouse; and, the dosage of the composition(s), concentration of components therein and timing of administering the composition(s), which elicit a suitable immunological response, such as by titrations of sera and analysis thereof for antibodies or antigens, e.g., by ELISA and/or RFFIT analysis. Such

determinations do not require undue experimentation from the knowledge of the skilled artisan, this disclosure and the documents cited herein. And, the time for sequential administrations may be ascertained without undue experimentation. For instance, dosages may be readily ascertained by those skilled in the art from this disclosure and the knowledge in the art. Thus, the skilled artisan may readily determine the amount of compound and optional additives, vehicles, and/or carrier in compositions and to be administered in methods of the invention. Typically, an adjuvant or additive is commonly used as 0.001 to 50 wt % solution in phosphate buffered saline, and the active ingredient is present in the order of micrograms to milligrams, such as about 0.0001 to about 5 wt %, preferably about 0.0001 to about 1 wt %, most preferably about 0.0001 to about 0.05 wt % or about 0.001 to about 20 wt %, preferably about 0.01 to about 10 wt %, and most preferably about 0.05 to about 5 wt %. Such determinations do not require undue experimentation from the knowledge of the skilled artisan, this disclosure and the documents cited herein. And, the time for sequential administrations may be ascertained without undue experimentation.

[00406] Examples of compositions which may comprise a therapeutic of the invention include liquid preparations for orifice, e.g., oral, nasal, anal, vaginal, peroral, intragastric, mucosal (e.g., perlingual, alveolar, gingival, olfactory or respiratory mucosa) etc., administration such as suspensions, syrups or elixirs; and, preparations for parenteral, subcutaneous, intradermal, intramuscular or intravenous administration (e.g., injectable administration), such as sterile suspensions or emulsions. Such compositions may be in admixture with a suitable carrier, diluent, or excipient such as sterile water, physiological saline, glucose or the like. The compositions may also be lyophilized. The compositions may contain auxiliary substances such as wetting or emulsifying agents, pH buffering agents, gelling or viscosity enhancing additives, preservatives, flavoring agents, colors, and the like, depending upon the route of administration and the preparation desired. Standard texts, such as "REMINGTON'S PHARMACEUTICAL SCIENCE", 17th edition, 1985, incorporated herein by reference, may be consulted to prepare suitable preparations, without undue experimentation.

[00407] Compositions of the invention, are conveniently provided as liquid preparations, e.g., isotonic aqueous solutions, suspensions, emulsions or viscous compositions which may be buffered to a selected pH. If digestive tract absorption is preferred, compositions of the invention may be in the "solid" form of pills, tablets, capsules, caplets and the like, including "solid" preparations which are time-released or which have a liquid filling, e.g., gelatin

covered liquid, whereby the gelatin is dissolved in the stomach for delivery to the gut. If nasal or respiratory (mucosal) administration is desired, compositions may be in a form and dispensed by a squeeze spray dispenser, pump dispenser or aerosol dispenser. Aerosols are usually under pressure by means of a hydrocarbon. Pump dispensers may preferably dispense a metered dose or, a dose having a particular particle size.

[00408] Compositions of the invention may contain pharmaceutically acceptable flavors and/or colors for rendering them more appealing, especially if they are administered orally. The viscous compositions may be in the form of gels, lotions, ointments, creams and the like (e.g., for transdermal administration) and will typically contain a sufficient amount of a thickening agent so that the viscosity is from about 2500 to 6500 cps, although more viscous compositions, even up to 10,000 cps may be employed. Viscous compositions have a viscosity preferably of 2500 to 5000 cps, since above that range they become more difficult to administer. However, above that range, the compositions may approach solid or gelatin forms, which are then easily administered as a swallowed pill for oral ingestion.

[00409] Liquid preparations are normally easier to prepare than gels, other viscous compositions, and solid compositions. Additionally, liquid compositions are somewhat more convenient to administer, especially by injection or orally. Viscous compositions, on the other hand, may be formulated within the appropriate viscosity range to provide longer contact periods with mucosa, such as the lining of the stomach or nasal mucosa.

[00410] Obviously, the choice of suitable carriers and other additives will depend on the exact route of administration and the nature of the particular dosage form, e.g., liquid dosage form (e.g., whether the composition is to be formulated into a solution, a suspension, gel or another liquid form), or solid dosage form (e.g., whether the composition is to be formulated into a pill, tablet, capsule, caplet, time release form or liquid-filled form).

[00411] Solutions, suspensions and gels, normally contain a major amount of water (preferably purified water) in addition to the active compound. Minor amounts of other ingredients such as pH adjusters (e.g., a base such as NaOH), emulsifiers or dispersing agents, buffering agents, preservatives, wetting agents, jelling agents, (e.g., methylcellulose), colors and/or flavors may also be present. The compositions may be isotonic, i.e., it may have the same osmotic pressure as blood and lacrimal fluid.

[00412] The desired isotonicity of the compositions of this invention may be accomplished using sodium chloride, or other pharmaceutically acceptable agents such as dextrose, boric

acid, sodium tartrate, propylene glycol or other inorganic or organic solutes. Sodium chloride is preferred particularly for buffers containing sodium ions.

[00413] Viscosity of the compositions may be maintained at the selected level using a pharmaceutically acceptable thickening agent. Methylcellulose is preferred because it is readily and economically available and is easy to work with. Other suitable thickening agents include, for example, xanthan gum, carboxymethyl cellulose, hydroxypropyl cellulose, carbomer, and the like. The preferred concentration of the thickener will depend upon the agent selected. The important point is to use an amount that will achieve the selected viscosity. Viscous compositions are normally prepared from solutions by the addition of such thickening agents.

[00414] A pharmaceutically acceptable preservative may be employed to increase the shelf-life of the compositions. Benzyl alcohol may be suitable, although a variety of preservatives including, for example, parabens, thimerosal, chlorobutanol, or benzalkonium chloride may also be employed. A suitable concentration of the preservative will be from 0.02% to 2% based on the total weight although there may be appreciable variation depending upon the agent selected.

[00415] Those skilled in the art will recognize that the components of the compositions should be selected to be chemically inert with respect to the active compound. This will present no problem to those skilled in chemical and pharmaceutical principles, or problems may be readily avoided by reference to standard texts or by simple experiments (not involving undue experimentation), from this disclosure and the documents cited herein.

[00416] It is generally envisaged that compounds and compositions of the invention will be administered by injection, as such compounds are to elicit anti-HIV antibodies, and the skilled artisan may, from this disclosure and the knowledge in the art, formulate compounds and compositions identified by herein methods for administration by injection and administer such compounds and compositions by injection.

[00417] The inventive compositions of this invention are prepared by mixing the ingredients following generally accepted procedures. For example the selected components may be simply mixed in a blender, or other standard device to produce a concentrated mixture which may then be adjusted to the final concentration and viscosity by the addition of water or thickening agent and possibly a buffer to control pH or an additional solute to control tonicity. Generally the pH may be from about 3 to 7.5. Compositions may be administered in dosages and by techniques well known to those skilled in the medical arts

taking into consideration such factors as the age, sex, weight, and condition of the particular patient, and the composition form used for administration (e.g., solid vs. liquid). Dosages for humans or other mammals may be determined without undue experimentation by the skilled artisan, from this disclosure, the documents cited herein, and the knowledge in the art.

[00418] Suitable regimes for initial administration and further doses or for sequential administrations also are variable, may include an initial administration followed by subsequent administrations; but nonetheless, may be ascertained by the skilled artisan, from this disclosure, the documents cited herein, and the knowledge in the art.

[00419] The present invention will be further illustrated in the following Examples which are given for illustration purposes only and are not intended to limit the invention in any way.

Examples

Example 1: HIV Vaccine Design to Target Germline Precursors of Glycan-Dependent Broadly Neutralizing Antibodies

[00420] See Steichen et al., Immunity Volume 45, Issue 3, 20 September 2016, Pages 483–496, the disclosure of which is incorporated by reference.

[00421] Broadly neutralizing antibodies (bnAbs) against the N332 supersite of the HIV Envelope (Env) trimer are the most common bnAbs induced during infection, making them promising leads for vaccine design. Wild-type Env glycoproteins lack detectable affinity for supersite-bnAb germline precursors and are therefore unsuitable immunogens to prime supersite-bnAb responses. Applicants employed mammalian cell surface display to design stabilized Env trimers with affinity for germline-reverted precursors of PGT121-class supersite bnAbs. The trimers maintained native-like antigenicity and structure, activated PGT121 inferred-germline B cells ex vivo when multimerized on liposomes, and primed PGT121-like responses in PGT121 inferred-germline knock-in mice. Design intermediates have levels of epitope modification between wild-type and germline-targeting trimers; their mutation gradient suggests sequential immunization to induce bnAbs, in which the germline-targeting prime is followed by progressively less-mutated design intermediates and lastly with native trimers. The vaccine design strategies described could be utilized to target other epitopes on HIV or other pathogens.

[00422] A vaccine is needed for global HIV prevention. Broadly neutralizing antibodies (bnAbs) directed against relatively conserved epitopes in the otherwise highly antigenically variable HIV Envelope (Env) glycoprotein trimer offer important guides for vaccine design. BnAbs have been isolated from a small minority of HIV-infected individuals and have been

shown to protect against challenge in various animal models, but have not been induced by vaccination in humans or standard animal models (Burton and Hangartner, 2016; Mascola and Haynes, 2013; West et al., 2014). BnAbs recovered from natural infection are typically highly mutated (Klein et al., 2013a; Mouquet et al., 2010; Pancera et al., 2010; Scheid et al., 2009; Walker et al., 2011; Xiao et al., 2009; Zhou et al., 2010) and many also contain insertions and/or deletions (Kepler et al., 2014), owing to chronic stimulation of B cells by mutating Env. Many bnAbs also possess unusually long or short heavy chain complementarity determining region 3 loops (Scheid et al., 2011; Walker et al., 2011; Walker et al., 2009; Wu et al., 2011; Zhou et al., 2010), and some are polyreactive (Haynes et al., 2005). Less mutated bnAbs with fewer unusual features have been engineered, offering more tractable goals for consistent vaccine elicitation (Georgiev et al., 2014; Jardine et al., 2016b; Sok et al., 2013). Overall, bnAb elicitation by vaccination presents a major challenge.

[00423] Recombinant native-like trimers are promising HIV vaccine components because they contain the conformational epitopes of most known bnAbs and lack many non-neutralizing epitopes present on less native constructs (Julien et al., 2013; Kong et al., 2016; Kwon et al., 2015; Lyumkis et al., 2013; Pancera et al., 2014; Sanders et al., 2013; Scharf et al., 2015). However, native-like trimers have features that may impede bnAb induction: they are highly glycosylated and expose both strain-specific neutralizing epitopes and non-neutralizing epitopes. Immunization with native-like trimers in standard mouse, rabbit and macaque models has thus far elicited either non-neutralizing antibodies (Hu et al., 2015) or neutralizing antibodies only against the immunogen strain (de Taeye et al., 2016; Sanders et al., 2015) analogous to the strain-specific responses to the seasonal flu vaccine in humans. Induction of HIV bnAbs will likely require development of vaccination strategies that focus responses to relatively conserved, sub-dominant epitopes and avoid or suppress responses to non-neutralizing and strain-specific epitopes.

[00424] Germline targeting, a vaccine priming strategy to initiate the affinity maturation of select germline-precursor B cells, could help solve this immunofocusing problem by preferentially activating bnAb precursors (Dimitrov, 2010; Xiao et al., 2009). The strategy aims to activate bnAb-precursor B cells, induce productive (bnAb-like) somatic mutations, and produce memory B cells that can be boosted subsequently to select additional productive mutations (Dosenovic et al., 2015; Jardine et al., 2015). For some bnAbs, inferred precursors have affinity for Env from particular HIV isolates (Andrabi et al., 2015; Doria-Rose et al., 2014; Gorman et al., 2016; Liao et al., 2013), facilitating design of priming immunogens

based on Env from those isolates (Haynes et al., 2012). For other bnAbs, efforts to identify wild-type Env that bind inferred precursors have failed (Hoot et al., 2013; Jardine et al., 2013; McGuire et al., 2013; Scheid et al., 2011; Xiao et al., 2009; Zhou et al., 2010). These latter cases require design of modified Env to serve as a priming immunogen (Dimitrov, 2010; Pancera et al., 2010; Xiao et al., 2009; Zhou et al., 2010). Proof of principle that designed germline-targeting immunogens can activate their intended precursors and generate a potentially boostable memory response was recently demonstrated in knock-in mice with B cell precursors for VRC01-class bnAbs directed to the CD4-binding site (Dosenovic et al., 2015; Jardine et al., 2015; McGuire et al., 2016). Following a germline-targeting prime, induction of bnAbs is expected to require a succession of boosts, driving a succession of germinal center reactions, in order to select sufficient mutations (Dimitrov, 2010; Dosenovic et al., 2015; Haynes et al., 2012; Jardine et al., 2013; 2015; 2016b; Klein et al., 2013b; Liao et al., 2013; McGuire et al., 2013; 2016; Pancera et al., 2010; Wu et al., 2011; Xiao et al., 2009; Zhou et al., 2010). Supporting the concept that sequential immunization with different immunogens will be required to develop a bnAb response, native-like trimers but not germline-targeting immunogens were found to boost near-bnAb B cells (bearing a mature VRC01-class bnAb heavy chain) to induce bnAbs (Dosenovic et al., 2015).

[00425] Glycan-dependent bnAbs in general, and N332-supersite bnAbs in particular, are important targets for germline-targeting vaccine design. In a recent longitudinal study of HIV infection in Africa, more than half of the HIV-infected individuals who produced bnAb responses produced them against glycan-directed epitopes, the majority of which were within the N332 supersite (Landais et al., 2016). The prevalence of N332-supersite bnAb responses is probably due in part to the high accessibility of their epitopes on top of the trimer.

[00426] Among N332-supersite bnAbs, PGT121-class bnAbs have been particularly well characterized, providing strong rationale for germline-targeting efforts. PGT121-class bnAbs are among the most potent bnAbs (Mouquet et al., 2012; Walker et al., 2011), and PGT121 delivered passively to macaques protects against SHIV infection (Moldt et al., 2012; Shingai et al., 2014) and can suppress viremia when delivered after infection (Barouch et al., 2013; Shingai et al., 2013). However, PGT121-class inferred precursors show no measureable affinity for wild-type Env proteins that have been evaluated (Mouquet et al., 2012; Sok et al., 2013). Thus, development of a priming immunogen for PGT121-class precursors requires either design of a modified Env or identification of a natural Env with PGT121-class germline-binding capacity. Crystal structures have been determined for several PGT121-

class bnAbs in complex with either BG505 SOSIP native-like trimers or gp120 (Garces et al., 2015; Garces et al., 2014; Julien et al., 2013; Kong et al., 2016; Pancera et al., 2014), and for unliganded structures of two germline-reverted PGT121 variants (Mouquet et al., 2012; Sok et al., 2013), providing critical information to guide design of modified Env for PGT121-class germline-targeting.

[00427] PGT121-class bnAbs interact with conformationally flexible structures on HIV Env, including several glycans and the V1 variable loop, making computational design of germline-targeting Env challenging. Here, Applicants developed a structure-guided directed evolution approach, using mammalian cell surface display, to design PGT121-class germline-targeting stabilized-trimer immunogens. Applicants multimerized these trimers on liposomes, and evaluated trimer and liposome immunogens by biophysical, structural, and ex vivo B cell activation analyses. Applicants further evaluated germline-targeting trimers by vaccination in PGT121 inferred-germline knock-in mice. Applicants' design process produced design intermediates with increasing levels of epitope modification between wild-type and germline-targeting trimers. These results led to Applicants' hypothesizing prime-boosting strategies in which a germline-targeting prime is followed by boosts with progressively less-modified design intermediates and then with wild-type Env followed ultimately by a cocktail of Env variants to expand breadth. Evaluation of several of these prime-boosting strategies in PGT121 germline and chimeric knock-in mice is described in a related study (Escolano et al., 2016).

[00428] Design of Germline-Targeting Gp120s. Applicants identified mammalian cell surface display as a desirable platform for engineering modified HIV Env constructs with affinity for inferred-germline PGT121 Abs, as it should allow for optimization of monomeric or multimeric antigens bearing mammalian glycans (Chen et al., 2008). Therefore, Applicants developed a lentivirus-based mammalian cell surface display method to carry out directed evolution of HIV gp120 monomers and gp140 trimers (Figure 9). Structural analysis of the PGT121 interaction with gp120 (Julien et al., 2013; Pancera et al., 2014) led us to hypothesize that the V1 and V3 loops were the key sites for germline-targeting mutations. For selection agents, Applicants assembled a collection of six germline-reverted Abs, all using heavy chain genes VH4-59, D3-3 and J6 and light chain genes V3-21 and J3, with varying degrees of mutation in the D gene and L-CDR3 and with differences in the non-templated regions at the V-D and D-J boundaries (Figure 10). Ranked by similarity to the germline D and L3 sequences, these Abs are: GL_{CDR3rev1} (most similar to germline D/L3),

GL_{CDR3rev2}, GL_{CDR3rev3}, GL_{CDR3rev4}, GL_{CDR3rev5}, and GL_{CDR3mat} (germline V and J genes but mature CDR3 loops). Applicants began by screening libraries based on BG505 T332N gp120 and BG505 SOSIP T332N gp140 (Sanders et al., 2013) in which the conserved glycosylation site at position 332 absent in BG505 was introduced (Figure 6A). These molecules had no detectable affinity for germline-reverted PGT121 Abs (Figure 6B and Table 1). Therefore, Applicants employed a “boot-strapping” approach: for initial screening Applicants utilized two variants of GL_{CDR3mat}, one with nine PGT121 light chain mutations (GL+9, with 28 μ M affinity for BG505 T332N gp120) and another with three light chain mutations (GL+3, no detectable affinity for BG505 T332N gp120) (Figure 10). Screening a gp120 random mutagenesis library for binding to GL+9 led to the molecule 3MUT, with mutations T135A and T139I to eliminate the V1 loop glycosylation sites at positions 133 and 137 (Figure 6A). Screening a gp140 structure-guided V1 loop library for binding to GL+3 led to the isolation of 5MUT, with four different mutations (V134Y, N136P, I138L, D140N) in the V1 loop. Combining the mutations in 3MUT and 5MUT produced 7MUT gp120, Applicants’ first construct with quantifiable affinity for GL_{CDR3mat} (K_D = 44 μ M, Figures 6A and 6B). To improve this affinity further, Applicants screened a gp120 V1 and V3 loop saturation mutagenesis library for binding to GL_{CDR3mat}; combining the most enriched mutations (N137F, T320F, Q328M) with 7MUT produced 10MUT, with $K_D \approx 1$ μ M for GL_{CDR3mat} (Figures 6A and 6B). Finally, to increase affinity and breadth, Applicants screened a gp120 V1 loop directed mutagenesis library for binding to GL_{CDR3rev2} and GL_{CDR3rev4} (Figure 10). This approach culminated in 11MUT_B, with K_D s of ~ 5 μ M, ~ 3 μ M and ~ 8 μ M for GL_{CDR3rev1}, GL_{CDR3rev3} and GL_{CDR3rev5}, respectively, and detectable but not quantifiable binding to GL_{CDR3rev4} (Figures 6A, 11, and Table 1). Thus, mammalian display-directed evolution enabled the design of germline-targeting gp120 molecules with appreciable affinity for PGT121 germline-reverted antibodies.

[00429] Design of Stabilized and Germline-Targeting Trimers. For initial design of germline-targeting and boosting trimers, Applicants transferred the germline-targeting mutations from the gp120 versions of 3MUT, 5MUT, 7MUT and 10MUT onto the BG505 SOSIP trimer platform. These molecules displayed characteristics of native-like trimers, such as high affinity for the trimer-specific bnAb PGT151 (Falkowska et al., 2014) and a melting temperature (T_m) similar to BG505 SOSIP (Figure 7A). Furthermore, all had similar monovalent affinities for PGT121 and GL_{CDR3mat} as their gp120 counterparts (Figure 7A), indicating that the germline-targeting mutations were transferable to a native-like trimer.

[00430] In addition to binding bnAb putative precursors, germline-targeting trimers should have an otherwise native-like antigenic profile, with high affinity for bnAbs and no significant affinity for non-neutralizing antibodies directed to epitopes exposed on monomeric gp120 but buried or conformationally absent on the trimer. BG505 SOSIP gp140, the trimer on which Applicants' PGT121-class germline-targeting designs were based, displays undesirable binding to V3 non-neutralizing antibodies (Sanders et al., 2013) (Figures 7B and 12) and induces non-neutralizing V3 responses in mice, rabbits, and macaques (de Taeye et al., 2016; Hu et al., 2015; Sanders et al., 2015), indicating that this trimer samples conformational states that expose non-neutralizing epitopes. Furthermore, BG505 SOSIP gp140 displayed on mammalian cells via a PDGFR linker showed strong binding to trimer-structure-dependent bnAbs (PGT151 and PGT145) (Falkowska et al., 2014; Walker et al., 2011) but also to non-neutralizing antibodies directed to the V3 loop (4025) (Gorny et al., 2011) and the CD4-binding site (b6) (Barbas et al., 1992) (not shown), suggesting the coexistence on the cell surface of native-like trimers along with non-native trimers, dimers and/or monomers. Applicants also found that adding germline-targeting mutations to BG505 SOSIP reduced the already-modest expression by 50% (Figure 7A). Therefore, Applicants sought to use mammalian display-directed evolution to improve the antigenic profile, thermal stability and expression level of the BG505 SOSIP trimer and germline-targeting trimers.

[00431] Applicants' trimer improvement effort focused on two types of libraries: (i) whole gene saturation mutagenesis libraries and (ii) a combinatorial library sampling the one or two most common HIV residues at Env positions where BG505 uses rare (frequency <10%) HIV residues (Figure 7C). The "rare" library, which allowed variation at eleven positions in gp120 and two in gp41, was screened for binding to trimer structure-preferring bnAbs PGT145, PGT151 and PG16 and for lack of binding to non-neutralizing antibodies b6 and 4025. This yielded the Rare3 clone with five mutations in gp120 (T106E, M271I, F288L, T290A, N363Q) and with the expression level improved by a factor of ~2 and the melting temperature (T_M) increased by 1.4 °C (Figure 13). The saturation mutagenesis library was constructed in three segments, two covering gp120 and one for gp41 (Figure 7C). Next generation sequencing and bioinformatics were employed to analyze the results of the first two sorts (Jardine et al., 2016a), while Sanger sequencing was used to identify enriched clones that survived four or five sorts. Enriched mutations from both sequencing methods were combined and tested in soluble trimers and were also assembled into combinatorial libraries and re-screened using the same antibodies as before. The gp41 library produced

MD2, with an L568D point mutation that increased expression levels by a factor of ~4, and MD33, with four additional mutations (F519S, A561P, V570H, R585H) that increased the T_M by 4 °C and improved expression by a factor of ~7 relative to BG505 SOSIP (Figure 13). The gp120 library produced MD16, with three mutations (F223W, R304V, A319Y) and reduced binding to V3 non-neutralizing antibodies (Figure 12). Finally, mutations from Rare3, MD16 and MD33 were combined to produce MD39 with 11 mutations (F223W and T290A did not improve the biophysical properties of the trimer and were excluded; data not shown). Compared to BG505 SOSIP.D664, the MD39 yield improved by a factor of ~7, T_M increased by °C, and antigenic profile improved by reduced V3 Ab reactivity and similar bnAb binding except for slightly reduced affinities for V2 apex bnAbs (Figures 7A, 7B and 13)

[00432] Combining the MD39 mutations with germline-targeting mutations produced germline-targeting trimers with improved properties. MD39-10MUT had 6-fold improved yield and 6°C higher T_M compared to 10MUT (Figures 7A and 13). Applicants' most advanced germline-targeting trimer, MD39-11MUT_B, the only trimer with detectable affinity for five of six PGT121-germline reverted variants tested (Figure 7B), had excellent yield, thermal stability and antigenic profile (Figure 7A,B). Directed evolution therefore produced native-like trimers with improved potential functionality via both stabilization and germline-targeting mutations.

[00433] Structural Analysis. To ascertain whether stabilized, PGT121-germline-targeting trimers maintain native-like structure, Applicants conducted crystallography and electron microscopy (EM) studies. Negative stain EM two-dimensional classification revealed that all four trimers tested (MD39, 10MUT, MD39-10MUT, and MD39-11MUT_B) were characterized by a high fraction ($\geq 95\%$) of native-like structural features and were similar in appearance to BG505 SOSIP. The MD39 mutations improved the structural uniformity of the 10MUT trimer, as the amount of flexible, native open conformations dropped from 35% to 5% between 10MUT and MD39-10MUT (see methods for description of 2D classification system). Applicants' best germline-targeting trimer, MD39-11MUT_B, exhibited 100% native closed conformations and was indistinguishable from BG505 SOSIP by EM. For higher resolution analysis, Applicants solved a 4.5 Å resolution crystal structure of MD39-10MUT_A, a variant of MD39-10MUT with one mutation added and another removed (Supplemental Methods), complexed with 35O22 and PGT124 (Garces et al., 2014; Sok et al., 2013). While this resolution precluded analyses of side-chain conformations, and the interface between

trimer and PGT124 could not be analyzed due to missing V1 loop density, the structure accurately determined the backbone positions for most (1659 of 1692) residues of gp140. Superposition of the gp140 backbones in this structure and in the 3.0 Å structure of BG505 SOSIP N137A complexed with 3H109L and 35O22 (PDBid: 5CEZ) or the 3.1 Å structure of BG505 SOSIP bound to PGT122 and 35O22 (PDBid: 4TVP) gives backbone rmsd values of 0.7 and 1.1 Å, respectively. Applicants conclude that MD39-10MUT_A, with 20 mutations relative to BG505 SOSIP T332N, retains an overall native-like conformation.

[00434] Sequential Boosting Strategies. As noted above, induction of bnAbs following a germline-targeting prime is expected to require sequential boosting with epitope variants to mature the response. With PGT121-class germline-targeting candidates (10MUT and 11MUT_B) in hand, Applicants developed boosting strategies aiming to select PGT121-like mutations and induce bnAbs. Applicants hypothesized that any sequential immunization strategy starting with a germline-targeting trimer should end with a native-like trimer, such as BG505 MD39 SOSIP, so as to select mutations productive for high-affinity interaction with the native trimer on the virus. However, in order for PGT121-class antibodies to engage their epitope including the N137 glycan on the V1 loop, such antibodies must accommodate V1 loops diverse not only in sequence but also in length and number of glycosylation sites (Figure 8A-C), implying that boosting with a single native-like trimer bearing a single V1 loop may not be sufficient. Indeed, boosting only with a BG505 native-like trimer would present a V1 loop that is significantly shorter than most (Figure 8B). Furthermore, modeling of variable loops and glycan conformations (not shown) suggested that diversity in the V2 and V4 loops might potentially impact the PGT121 epitope by altering conformational sampling of the V1 loop or N332 glycan, respectively (Figure 8A-C), and immunodominant responses involving V2 or V4 could potentially sterically interfere with PGT121-class boosting. Based on these considerations, Applicants hypothesized that a cocktail of native-like trimers displaying diverse variants of the V1, V2, and V4 loops, especially variants within “hotspots” of more frequently-occurring combinations of length and number of glycosylation sites (Figure 8B), might be needed to select PGT121-class mutations favoring neutralization breadth. Applicants therefore designed and produced four native-like trimers based on BG505 MD39 SOSIP containing diverse loops for V1, V2, V4, and V5 (Figures 8C, 14A and 14B). These trimers, together with BG505 MD39 SOSIP, form a five-member variable loop cocktail (VLC) that might broaden PGT121-like responses initiated by a germline-targeting trimer (Figure 8D).

[00435] Applicants then considered the question of what intermediate boosts, if any, might be employed between a germline-targeting prime and a native-like trimer. Applicants' germline-targeting design intermediates become increasingly more native-like in the PGT121 epitope (e.g. 7MUT, 5MUT and 3MUT have six, four and two epitope mutations, respectively), but the 5MUT and 3MUT mutations are mutually exclusive (3MUT lacks two V1 glycans while 5MUT has those glycans but has four other V1 loop mutations) (Figure 6A). These considerations impose directionality on any boosting scheme (e.g. 7MUT should not be used after 5MUT or 3MUT or WT, and 3MUT should not be used after 5MUT), thus limiting the number of possible schemes (Table 2). Considering only the most efficient directional schemes, those employing boosting pairs that differ by more than one mutation or involve substantial affinity changes (Table 3), Applicants identified a total of seven potential boosting schemes (Figure 8E).

[00436] Applicants sought to rank these schemes to allow prioritization for experimental testing. Applicants reasoned that the least mutated antibody that shows measurable affinity for all of the potential boosting immunogens, GL+9, could serve as a proxy for intermediate PGT121-class antibodies developing after a germline-targeting prime and before a native-like boost. Applicants further reasoned that the affinity drop, the ratio of GL+9 affinities for two immunogens, could be used to estimate the likelihood of successfully boosting memory B cells when the two immunogens are used in sequence (e.g. the GL+9 K_{DS} for 7MUT and 3MUT are 3 nM and 1600 nM respectively, so when immunizing with 7MUT followed by 3MUT, the affinity drop would be $1600/3=530$). One expects that a boost immunogen with very different epitope structure from the previous immunogen may result in too large an affinity drop to activate memory B cells generated by the prior immunogen. Applicants estimated the affinity drops for all seven boosting schemes (Figure 8E), ranked them according to the largest affinity drop in that boosting scheme, and listed the three most likely to succeed (Figure 8F).

[00437] In collaborating work, Escolano and colleagues (Escolano *et al.*, 2016) evaluated boosting schemes following the 10MUT trimer prime in PGT121 germline (GL_{CDR3-rev4}) knock-in mice and PGT121 mature-heavy-chain/germline-light-chain knock-in mice. Relying on the directionality of the boosting immunogens developed here, Escolano *et al.* used serum ELISA against boost candidates after each immunization to select the most native-like directional boost for which at least weak serum reactivity could be detected; that process resulted in the testing of the second scheme in Figure 8F and the finding that this

scheme induces PGT121-like bnAbs with substantial breadth and potency. While the first scheme in 12F remains to be tested, the data in Escolano et al. support the validity of the logic underlying these boosting schemes.

[00438] Applicants note that the affinity drop analysis also provides clues on how to improve boosting schemes: to minimize the probability of a boost failure at a high affinity drop, one could redesign immunogens to equalize the affinity drops in any given scheme. Thus the germline-targeting design process is capable of defining potential boost immunogens and directional boosting schemes, and it can guide prioritization and improvement of such schemes.

[00439] Germline-targeting vaccine design offers the potential to initiate the induction of specific classes of protective antibodies against HIV or other pathogens that have eluded vaccine development. Many protective bnAbs against HIV are directed toward glycan-dependent epitopes on the trimeric glycoprotein spike (Burton and Hangartner, 2016). Therefore, methods are needed to develop trimer immunogens for germline targeting and boosting of glycan-dependent bnAbs. Trimer immunogens should be stabilized, to maximize the probability of retaining native-like conformational epitopes in vivo and to minimize the probability of eliciting non-neutralizing Abs that could potentially detract from priming or boosting the targeted bnAb responses.

[00440] Here, Applicants (1) developed a mammalian cell surface display directed evolution method for optimization of multimeric antigens bearing human glycans; (2) engineered stabilized HIV Env trimers with affinity for both germline and mature PGT121-class glycan-dependent bnAbs; (3) showed by crystallography and electron microscopy that these trimers maintain native-like conformations; (4) demonstrated that germline-targeting trimers multimerized on liposomes potently activate PGT121 germline and mature B cells ex vivo; and (5) showed that soluble germline-targeting trimers can prime PGT121-class responses in vivo, in a PGT121 inferred-germline knock-in mouse. Applicants' data indicate that 11MUT_B trimers and trimer-liposomes are promising candidates for priming PGT121-class glycan-dependent bnAb responses in immune systems with diverse antibody repertoires, although the frequency of PGT121-class precursors in humans and the germline-targeting affinities/avidities necessary to prime those precursors remain to be determined.

[00441] This work may provide a more general template for HIV bnAb germline-targeting compared to previous work on germline-targeting for VRC01-class bnAbs directed to the

CD4-binding site. VRC01-class bnAbs generally do not depend on glycans for their activity, evidenced by the fact that elimination of glycans surrounding the VRC01 epitope generally increases neutralization potency (Jardine et al., 2016b); this has led to removal of all epitope-proximal glycans from germline-targeting candidates (Jardine et al., 2013; 2015; 2016a; McGuire et al., 2013; 2014; 2016). However, the activity of many HIV bnAbs requires engagement of one or more glycans within their epitope, and germline-targeting primes should probably retain such key glycans, as was the case here with the N332 glycan. Furthermore, owing to the relative inaccessibility of the VRC01 epitope on native-like trimers, efforts to design VRC01-class germline-targeting primes have converged on strategies to increase epitope exposure by presentation on minimal domains rather than on trimers (Jardine et al., 2013; 2015; 2016a; McGuire et al., 2013; 2014; 2016), although boosting with native-like trimers is anticipated to be required to mature the response (Jardine et al., 2016b). In contrast, many proteo-glycan epitopes are well exposed on native-like trimers, and some are formed only on intact trimers, making native-like trimers like those designed here the preferred platform for germline-targeting. Indeed, multiple different bnAbs could potentially be primed with a single trimer harboring multiple germline-targeting epitopes.

[00442] As germline-targeting vaccine design requires developing not only the vaccine prime but also boost immunogens to mature the response in order to elicit bnAbs, Applicants developed both a stabilized native-like trimer (MD39) and a cocktail of native-like trimers (VLC cocktail) that could be employed as boosts to refine and expand the breadth of responses initiated by a germline-targeting prime. However, considering that memory B cells induced by the germline-targeting prime may not be sufficiently mutated to be boosted by a native-like trimer, intermediate boosts may be needed to mature the response prior to native-like boosts. In the process of developing PGT121-class germline-targeting immunogens, Applicants created design intermediates with increasing levels of epitope modification between wild-type and germline-targeting trimers. These molecules are candidate boost immunogens that if used in sequence offer directional and gradual epitope changes to guide maturation of the B cell response. Seven potential sequential immunization schemes were proposed, and Applicants' analysis of affinity drops provided a ranking of those schemes. In a related paper (Escolano et al., 2016), a subset of these prime-boosting schemes were evaluated in PGT121 germline knock-in mice and PGT121 mature-heavy-chain/germline-light-chain knock-in mice, and one such scheme was found to be effective for inducing

bnAbs in both mouse models, supporting the germline-targeting vaccine design process described here and encouraging its expanded use and further improvement.

[00443] While here Applicants have described strategies for designing trimer immunogens with changes in the structure of an epitope in order to prime and mature an epitope-specific response, ultimate success of these strategies may also require modification of antigenic surfaces outside the epitope of interest, to minimize boosting of off-target responses that might hinder or interfere with the desired epitope-specific response.

[00444] The approaches employed here could be adapted for immunogen design to other bnAb targets on HIV and other pathogens. The “bootstrapping” strategy of using partially mutated antibodies (such as GL+3 and GL+9) as initial selection agents and then using antibodies closer to germline in successive iterations, could be useful for design of germline-targeting and boosting immunogens for other bnAbs, such as HIV V2 Apex glycan-dependent bnAbs (Andrabi et al., 2015; Gorman et al., 2016) or influenza virus hemagglutinin stem-directed bnAbs. Applicants’ mammalian display methods allowing directed evolution on native-like trimers should be useful in those endeavors and could also be used to stabilize monomeric or multimeric glycoprotein immunogens for diverse viral vaccines.

[00445] In summary, Applicants have developed stabilized native-like trimer immunogens for germline-targeting and boosting of glycan-dependent PGT121-class bnAb responses against HIV. The immunogens and boosting schemes Applicants created are candidates for human vaccine testing and further optimization, while the methods developed here are applicable to immunogen design for other epitopes and pathogens and thus are of relevance for future vaccine design.

[00446] DNA gene synthesis and protein production. Genes were synthesized by GenScript. Gp120s, gp140s, Fabs and IgGs were expressed in 293 cells and purified as described below.

[00447] Library assembly. The BG505 SOSIP whole gene saturation mutagenesis and “rare amino acid” libraries were synthesized by Integrated DNA Technologies and GenScript, respectively. Libraries for germline targeting were created by error prone PCR (gene morph II Agilent), site directed mutagenesis (QuikChange Agilent) or 2-step assembly PCR of degenerate primers using the Q5 High-Fidelity DNA Polymerase (New England Biolabs) and cloned into a modified version of the gateway cloning entry vector pENTR/D-TOPO (Ota et al., 2012) using the circularpolymeraseextension cloning (CPEC) method (Quan and

Tian, 2014) or Gibson Assembly (New England Biolabs) according to manufacturer's instructions. All libraries were then transferred to the lentiviral vector pLenti CMVTRE3G puro Dest (Ota et al., 2012) using the LR Clonase II enzyme mix (Thermo Scientific).

[00448] Lentivirus production and stable cell generation. 293T cells cultured in Advanced DMEM (Gibco) supplemented with 5% FCS, GlutaMAX (Gibco), 2-mercaptoethanol (Gibco) and Antibiotic-Antimycotic (Gibco) were co-transfected with 10.8 µg pLenti CMVTRE3G puro Dest gene library, 7.0 µg psPAX2 and 3.8 µg pMD2.G as described (Salmon and Trono, 2007). 293T cells stably expressing rtTA3G from the pLenti CMV rtTA3G Blast vector (Ota et al., 2012) were transduced at low moi (<0.1) in a T75 or T225 flask in the presence of 10 µg/mL blasticidin and after 24h transferred to medium supplemented with 2 µg/mL puromycin.

[00449] Cell surface expression and FACS. 293T cells containing the stable library were induced with doxycycline (1µg/mL) and harvested the next day in FACS buffer (HBSS, 1 mM EDTA, 0.5% BSA). Cells containing BG505-SOSIP libraries were transfected with furin 24 h prior to induction. Cells were stained with IgGs or Fabs for ~30 min, washed with FACS buffer, and then stained with fluorescein isothiocyanate (FITC)-labeled α-cMyc (Immunology Consultants Laboratory). IgGs were labeled with phycoerythrin (PE)-conjugated α-human IgG (Sigma), Fabs containing HA epitope tags (PGT145, PGT151, and PG16) were labeled with α-HA-PE (Miltenyi Biotec) and Fabs containing V5 epitope tags (B6 and 4025) were labeled with α-V5-FITC (GeneTex). Cells were sorted on a BD Influx (BD Biosciences) FACS sorter. Approximately 2x10⁵ double positive cells were collected and expanded for ~one week in the presence of puromycin and blasticidin before the next round of enrichment. Once the desired population had been obtained, chromosomal DNA was extracted from the cell culture using the GenElute Mammalian Genomic DNA Miniprep Kit (Sigma). The gp120 or gp140 gene was PCR amplified from the genomic DNA and inserted back into the pENTR vector using CPEC cloning or Gibson assembly, transformed into top10 competent cells (Invitrogen) and colonies were sequenced at Genewiz.

[00450] Next generation sequencing (NGS). Sequencing and bioinformatic analysis of the BG505-SOSIP whole gene saturation mutagenesis libraries were done essentially as described previously (Jardine et al., 2016a).

[00451] Trimer-conjugated liposome synthesis and characterization. Unilamellar liposomes comprised of DSPC:cholesterol:DGS-NTA(Ni) lipids in a 66.5:28.5:5 mole ratio were synthesized by lipid film rehydration and membrane extrusion, followed by post-

synthesis binding of 6xHis-tagged trimer for 2 hrs at 4 °C. Unconjugated trimer was removed by size exclusion chromatography. Total conjugated trimer was quantified by ELISA in the presence of 1% triton-X and 100 mM imidazole to fully disrupt liposomes and Ni-6xHis interactions, respectively, for uninhibited detection via an α -6xHis antibody. Antigenic profiles were determined by ELISA on intact liposomes.

[00452] Ca^{2+} -flux measurements and Immunizations. Details about Ca^{2+} -flux assays and mouse immunizations can be found in Extended experimental procedures and in (Escolano et al., 2016).

[00453] Negative-stain electron microscopy. Purified SOSIP trimers were analyzed by negative stain EM using a protocol adapted from (de Taeye et al., 2016).

[00454] Differential scanning calorimetry (DSC) and Surface plasmon resonance (SPR) methods are described in the Extended experimental procedures.

[00455] Table 1. The binding affinities of germline targeting gp120s, related to FIG. 6.

BG505- gp120	PGT121	3H3L	GL+9	GL+3	GL _{CDR3} - RM1	GL _{CDR3} - rev5	GL _{CDR3} - rev4	GL _{CDR3} - rev3	GL _{CDR3} - rev2	GL _{CDR3} - rev1	GL _{CDR3} -GL _{H-rev4} 121 _L	121 _H GL _{L-rev4}
WT (T332N)	7.5	250	28000	>128000	>128000	>8000	>84000	>128000	-	>8000	600	>38000
2MUT	2.7	-	4900	>40000	-	-	>40000	-	-	-	63	>40000
3MUT	4.6	19	1600	>28000	>21000	-	>28000	>11000	-	-	22	>28000
5MUT	5.7	2.5	18	WB	WB	-	>34000	-	-	-	6	13000
6MUT	1.4	-	19	>24000	-	-	>24000	-	-	-	5	>24000
7MUT	1.2	0.25	3	12200	44000	-	>36000	-	-	-	1.3	>36000
9MUT _A	0.57	-	-	2700	2900	-	>70000	-	-	-	-	57000
9MUT _B	1.5	28	-	29000	WB	-	>107000	-	-	-	220	39000
10MUT	0.59	0.04	1.2	1200	790	WB	>150000	WB	>150000	WB	-	47000
10MUT-KO	435	-	-	-	>21000	-	>21000	-	-	-	20000	>21000
11MUT _A	-	-	-	-	1200	-	WB	-	-	-	-	51000
11MUT _B	0.15	0.075	0.6	600	840	7700	WB	3000	-	5200	-	-

Values are K_{DS} (nM) measured by SPR.

WB, weak binding, not quantified.

-, not measured.

[00456]

[00457] Table 2. Sequential boosting pairs that were eliminated based on violation of directionality, related to FIG. 8.

Sequential boosting pair	Directionality violation
11B → 10/9A	11B contains the native residue N137 which is mutated to F in 10/9A
6 → 3	6 contains the native glycosylation site at N133 which is mutated in 3
5 → 3	5 contains native glycosylation sites at N133 and N137 and both are mutated in 3
5 → 2	5 contains native glycosylation site at N137 which is mutated in 2

Any boosting pair in which the first immunogen contains a native residue that is mutated in the second immunogen is a violation of directionality. The immunogen names have "MUT" removed for simplicity.

[00458] Table 3. Characteristics of sequential boosting pairs that obey directionality, related to FIG. 8.

Sequential boosting pair	Affinity drop #	# of AA changes	# of AA closer to WT	comment
11B/10 → 7	Small (5/3)	6/3	4/3	Shown in Figure 8.
11B/10 → 6	Medium (32/16)	7/4	5/4	Shown in Figure 8.
11B/10 → 5	Medium (30/15)	7/5	6/5	Shown in Figure 8.
11B/10 → 3	Large (2700/1300)	9/7	8/7	Shown in Figure 8.
11B/10 → WT	Large (47000/23000)	10/9	10/9	Shown in Figure 8.
7 → 5	Small (6)	2	2	Shown in Figure 8.
7 → 3	Medium (530)	4	4	Shown in Figure 8.
7 → WT	Large (9300)	6	6	Shown in Figure 8.
6 → WT	Large (1500)	5	5	Shown in Figure 8.
5 → WT	Large (1600)	4	4	Shown in Figure 8.
3 → WT	Medium (18)	2	2	Shown in Figure 8.
11B/10 → 2	Large (8200/4100)	10/8	9/8	Would be followed by: 2 → WT
10 → 9A	Small (4)*	1	1	Small affinity drop and only 1 mutation, thus inefficient
7 → 8	Small (5)	1	1	Small affinity drop and only 1 mutation, thus inefficient
6 → 5	Small (1)	1	1	Small affinity drop and only 1 mutation, thus inefficient
6 → 2	Medium (260)	4	4	Would be followed by: 2 → WT
3 → 2	Small (3)	1	1	Small affinity drop and only 1 mutation, thus inefficient
2 → WT	Small (5)	1	1	Small affinity drop and only 1 mutation, thus inefficient

#, Affinity drops were calculated based on binding to the GL+9 antibody, as described in the text, except where noted otherwise. Affinity drops were defined as small (<10), medium (10-1000), or large (>1000).

*, Affinity drops were calculated based on binding to the GL_{CDR3-mat} antibody.

The immunogen names have "MUT" removed for simplicity.

WT, BG505-T332N.

References

- Andrabi, R., Voss, J.E., Liang, C.H., Briney, B., McCoy, L.E., Wu, C.Y., Wong, C.H., Poignard, P., and Burton, D.R. (2015). Identification of Common Features in Prototype Broadly Neutralizing Antibodies to HIV Envelope V2 Apex to Facilitate Vaccine Design. *Immunity* 43, 959-973.
- Barbas, C.F., 3rd, Bjorling, E., Chiodi, F., Dunlop, N., Cababa, D., Jones, T.M., Zebedee, S.L., Persson, M.A., Nara, P.L., Norrby, E., *et al.* (1992). Recombinant human Fab fragments neutralize human type 1 immunodeficiency virus in vitro. *Proceedings of the National Academy of Sciences of the United States of America* 89, 9339-9343.
- Barouch, D.H., Whitney, J.B., Moldt, B., Klein, F., Oliveira, T.Y., Liu, J., Stephenson, K.E., Chang, H.W., Shekhar, K., Gupta, S., *et al.* (2013). Therapeutic efficacy of potent neutralizing HIV-1-specific monoclonal antibodies in SHIV-infected rhesus monkeys. *Nature* 503, 224-228.
- Burton, D.R., and Hangartner, L. (2016). Broadly Neutralizing Antibodies to HIV and Their Role in Vaccine Design. *Annu Rev Immunol* 34, 635-659.
- Chen, K.C., Wu, C.H., Chang, C.Y., Lu, W.C., Tseng, Q., Prijovich, Z.M., Schechinger, W., Liaw, Y.C., Leu, Y.L., and Roffler, S.R. (2008). Directed evolution of a lysosomal enzyme with enhanced activity at neutral pH by mammalian cell-surface display. *Chem Biol* 15, 1277-1286.
- de Taeye, S.W., Moore, J.P., and Sanders, R.W. (2016). HIV-1 Envelope Trimer Design and Immunization Strategies To Induce Broadly Neutralizing Antibodies. *Trends Immunol* 37, 221-232.
- Dimitrov, D.S. (2010). Therapeutic antibodies, vaccines and antibodyomes. *mAbs* 2, 347-356.
- Doria-Rose, N.A., Schramm, C.A., Gorman, J., Moore, P.L., Bhiman, J.N., DeKosky, B.J., Ernandes, M.J., Georgiev, I.S., Kim, H.J., Pancera, M., *et al.* (2014). Developmental pathway for potent V1V2-directed HIV-neutralizing antibodies. *Nature* 509, 55-62.
- Dosenovic, P., von Boehmer, L., Escolano, A., Jardine, J., Freund, N.T., Gitlin, A.D., McGuire, A.T., Kulp, D.W., Oliveira, T., Scharf, L., *et al.* (2015). Immunization for HIV-1 Broadly Neutralizing Antibodies in Human Ig Knockin Mice. *Cell* 161, 1505-1515.
- Escolano, A., Steichen, J., Dosenovic, P., Kulp, D.W., Golijanin, J., Sok, D., Freund, N.T., Gitlin, A.D., Oliveira, T., Araki, T., *et al.* (in-press). Sequential Immunization Elicits Broadly Neutralizing anti-HIV-1 Antibodies in Ig Knock-in Mice. *Cell*.
- Falkowska, E., Le, K.M., Ramos, A., Doores, K.J., Lee, J.H., Blattner, C., Ramirez, A., Derking, R., van Gils, M.J., Liang, C.H., *et al.* (2014). Broadly neutralizing HIV antibodies define a glycan-dependent epitope on the prefusion conformation of gp41 on cleaved envelope trimers. *Immunity* 40, 657-668.

- Garces, F., Lee, J.H., de Val, N., de la Pena, A.T., Kong, L., Puchades, C., Hua, Y., Stanfield, R.L., Burton, D.R., Moore, J.P., *et al.* (2015). Affinity Maturation of a Potent Family of HIV Antibodies Is Primarily Focused on Accommodating or Avoiding Glycans. *Immunity* 43, 1053-1063.
- Garces, F., Sok, D., Kong, L., McBride, R., Kim, H.J., Saye-Francisco, K.F., Julien, J.P., Hua, Y., Cupo, A., Moore, J.P., *et al.* (2014). Structural evolution of glycan recognition by a family of potent HIV antibodies. *Cell* 159, 69-79.
- Georgiev, I.S., Rudicell, R.S., Saunders, K.O., Shi, W., Kirys, T., McKee, K., O'Dell, S., Chuang, G.Y., Yang, Z.Y., Ofek, G., *et al.* (2014). Antibodies VRC01 and 10E8 neutralize HIV-1 with high breadth and potency even with Ig-framework regions substantially reverted to germline. *Journal of immunology* 192, 1100-1106.
- Gorman, J., Soto, C., Yang, M.M., Davenport, T.M., Guttman, M., Bailer, R.T., Chambers, M., Chuang, G.Y., DeKosky, B.J., Doria-Rose, N.A., *et al.* (2016). Structures of HIV-1 Env V1V2 with broadly neutralizing antibodies reveal commonalities that enable vaccine design. *Nat Struct Mol Biol* 23, 81-90.
- Gorny, M.K., Sampson, J., Li, H., Jiang, X., Totrov, M., Wang, X.H., Williams, C., O'Neal, T., Volsky, B., Li, L., *et al.* (2011). Human anti-V3 HIV-1 monoclonal antibodies encoded by the VH5-51/VL lambda genes define a conserved antigenic structure. *PLoS One* 6, e27780.
- Haynes, B.F., Fleming, J., St Clair, E.W., Katinger, H., Stiegler, G., Kunert, R., Robinson, J., Scearce, R.M., Plonk, K., Staats, H.F., *et al.* (2005). Cardiolipin polyspecific autoreactivity in two broadly neutralizing HIV-1 antibodies. *Science* 308, 1906-1908.
- Haynes, B.F., Kelsoe, G., Harrison, S.C., and Kepler, T.B. (2012). B-cell-lineage immunogen design in vaccine development with HIV-1 as a case study. *Nature biotechnology* 30, 423-433.
- Hoot, S., McGuire, A.T., Cohen, K.W., Strong, R.K., Hangartner, L., Klein, F., Diskin, R., Scheid, J.F., Sather, D.N., Burton, D.R., *et al.* (2013). Recombinant HIV envelope proteins fail to engage germline versions of anti-CD4bs bNAbs. *PLoS pathogens* 9, e1003106.
- Hu, J.K., Crampton, J.C., Cupo, A., Ketas, T., van Gils, M.J., Sliepen, K., de Taeye, S.W., Sok, D., Ozorowski, G., Deresa, I., *et al.* (2015). Murine Antibody Responses to Cleaved Soluble HIV-1 Envelope Trimers Are Highly Restricted in Specificity. *J Virol* 89, 10383-10398.
- Jardine, J., Julien, J.P., Menis, S., Ota, T., Kalyuzhniy, O., McGuire, A., Sok, D., Huang, P.S., MacPherson, S., Jones, M., *et al.* (2013). Rational HIV immunogen design to target specific germline B cell receptors. *Science* 340, 711-716.
- Jardine, J., Sok, D., Julien, J.P., Briney, B., Sarkar, A., Adachi, Y., Dewanji, D., Hsueh, J., Jones, M., Kalyuzhniy, I., *et al.* (2016b). Minimally Mutated HIV-1 Broadly Neutralizing Antibodies to Guide Reductionist Vaccine Design. *PLoS Pathogens*.

- Jardine, J.G., Kulp, D.W., Havenar-Daughton, C., Sarkar, A., Briney, B., Sok, D., Sesterhenn, F., Ereno-Orbea, J., Kalyuzhniy, O., Deresa, I., et al. (2016a). HIV-1 broadly neutralizing antibody precursor B cells revealed by germline-targeting immunogen. *Science* 351, 1458-1463.
- Jardine, J.G., Ota, T., Sok, D., Pauthner, M., Kulp, D.W., Kalyuzhniy, O., Skog, P.D., Thinnies, T.C., Bhullar, D., Briney, B., et al. (2015). Priming a broadly neutralizing antibody response to HIV-1 using a germline-targeting immunogen. *Science* 349, 156-161.
- Julien, J.P., Cupo, A., Sok, D., Stanfield, R.L., Lyumkis, D., Deller, M.C., Klasse, P.J., Burton, D.R., Sanders, R.W., Moore, J.P., et al. (2013). Crystal structure of a soluble cleaved HIV-1 envelope trimer. *Science* 342, 1477-1483.
- Kepler, T.B., Liao, H.X., Alam, S.M., Bhaskarabhatla, R., Zhang, R., Yandava, C., Stewart, S., Anasti, K., Kelsoe, G., Parks, R., et al. (2014). Immunoglobulin gene insertions and deletions in the affinity maturation of HIV-1 broadly reactive neutralizing antibodies. *Cell Host Microbe* 16, 304-313.
- Klein, F., Diskin, R., Scheid, J.F., Gaebler, C., Mouquet, H., Georgiev, I.S., Pancera, M., Zhou, T., Incesu, R.B., Fu, B.Z., et al. (2013a). Somatic mutations of the immunoglobulin framework are generally required for broad and potent HIV-1 neutralization. *Cell* 153, 126-138.
- Klein, F., Diskin, R., Scheid, J.F., Gaebler, C., Mouquet, H., Georgiev, I.S., Pancera, M., Zhou, T., Incesu, R.B., Fu, B.Z., et al. (2013a). Somatic mutations of the immunoglobulin framework are generally required for broad and potent HIV-1 neutralization. *Cell* 153, 126-138.
- Klein, F., Mouquet, H., Dosenovic, P., Scheid, J.F., Scharf, L., and Nussenzweig, M.C. (2013b). Antibodies in HIV-1 vaccine development and therapy. *Science* 341, 1199-1204.
- Kong, R., Xu, K., Zhou, T., Acharya, P., Lemmin, T., Liu, K., Ozorowski, G., Soto, C., Taft, J.D., Bailer, R.T., et al. (2016). Fusion peptide of HIV-1 as a site of vulnerability to neutralizing antibody. *Science* 352, 828-833.
- Kwon, Y.D., Pancera, M., Acharya, P., Georgiev, I.S., Crooks, E.T., Gorman, J., Joyce, M.G., Guttman, M., Ma, X., Narpala, S., et al. (2015). Crystal structure, conformational fixation and entry-related interactions of mature ligand-free HIV-1 Env. *Nat Struct Mol Biol* 22, 522-531.
- Landais, E., Huang, X., Havenar-Daughton, C., Murrell, B., Price, M.A., Wickramasinghe, L., Ramos, A., Bian, C.B., Simek, M., Allen, S., et al. (2016). Broadly Neutralizing Antibody Responses in a Large Longitudinal Sub-Saharan HIV Primary Infection Cohort. *PLoS pathogens* 12, e1005369.
- Liao, H.X., Lynch, R., Zhou, T., Gao, F., Alam, S.M., Boyd, S.D., Fire, A.Z., Roskin, K.M., Schramm, C.A., Zhang, Z., et al. (2013). Co-evolution of a broadly neutralizing HIV-1 antibody and founder virus. *Nature* 496, 469-476.

- Lyumkis, D., Julien, J.P., de Val, N., Cupo, A., Potter, C.S., Klasse, P.J., Burton, D.R., Sanders, R.W., Moore, J.P., Carragher, B., *et al.* (2013). Cryo-EM structure of a fully glycosylated soluble cleaved HIV-1 envelope trimer. *Science* 342, 1484-1490.
- Mascola, J.R., and Haynes, B.F. (2013). HIV-1 neutralizing antibodies: understanding nature's pathways. *Immunological reviews* 254, 225-244.
- McGuire, A.T., Hoot, S., Dreyer, A.M., Lippy, A., Stuart, A., Cohen, K.W., Jardine, J., Menis, S., Scheid, J.F., West, A.P., *et al.* (2013). Engineering HIV envelope protein to activate germline B cell receptors of broadly neutralizing anti-CD4 binding site antibodies. *The Journal of experimental medicine* 210, 655-663.
- McGuire, A.T., Gray, M.D., Dosenovic, P., Gitlin, A.D., Freund, N.T., Petersen, J., Correnti, C., Johnsen, W., Kegel, R., Stuart, A.B., *et al.* (2016). Specifically modified Env immunogens activate B-cell precursors of broadly neutralizing HIV-1 antibodies in transgenic mice. *Nat Commun* 7, 10618.
- Moldt, B., Rakasz, E.G., Schultz, N., Chan-Hui, P.Y., Swiderek, K., Weisgrau, K.L., Piaskowski, S.M., Bergman, Z., Watkins, D.I., Poignard, P., *et al.* (2012). Highly potent HIV-specific antibody neutralization in vitro translates into effective protection against mucosal SHIV challenge in vivo. *Proceedings of the National Academy of Sciences of the United States of America* 109, 18921-18925.
- Mouquet, H., Scharf, L., Euler, Z., Liu, Y., Eden, C., Scheid, J.F., Halper-Stromberg, A., Gnanapragasam, P.N., Spencer, D.I., Seaman, M.S., *et al.* (2012). Complex-type N-glycan recognition by potent broadly neutralizing HIV antibodies. *Proceedings of the National Academy of Sciences of the United States of America* 109, E3268-3277.
- Mouquet, H., Scheid, J.F., Zoller, M.J., Krogsgaard, M., Ott, R.G., Shukair, S., Artyomov, M.N., Pietzsch, J., Connors, M., Pereyra, F., *et al.* (2010). Polyreactivity increases the apparent affinity of anti-HIV antibodies by heterologation. *Nature* 467, 591-595.
- Ota, T., Doyle-Cooper, C., Cooper, A.B., Huber, M., Falkowska, E., Doores, K.J., Hangartner, L., Le, K., Sok, D., Jardine, J., *et al.* (2012). Anti-HIV B Cell lines as candidate vaccine biosensors. *Journal of immunology* 189, 4816-4824.
- Pancera, M., McLellan, J.S., Wu, X., Zhu, J., Changela, A., Schmidt, S.D., Yang, Y., Zhou, T., Phogat, S., Mascola, J.R., *et al.* (2010). Crystal structure of PG16 and chimeric dissection with somatically related PG9: structure-function analysis of two quaternary-specific antibodies that effectively neutralize HIV-1. *J Virol* 84, 8098-8110.
- Pancera, M., Zhou, T., Druz, A., Georgiev, I.S., Soto, C., Gorman, J., Huang, J., Acharya, P., Chuang, G.Y., Ofek, G., *et al.* (2014). Structure and immune recognition of trimeric pre-fusion HIV-1 Env. *Nature* 514, 455-461.
- Pugach, P., Ozorowski, G., Cupo, A., Ringe, R., Yasmeen, A., de Val, N., Derking, R., Kim, H.J., Korzun, J., Golabek, M., *et al.* (2015). A native-like SOSIP.664 trimer based on an HIV-1 subtype B env gene. *J Virol* 89, 3380-3395.

- Quan, J., and Tian, J. (2014). Circular polymerase extension cloning. *Methods in molecular biology* 1116, 103-117.
- Salmon, P., and Trono, D. (2007). Production and titration of lentiviral vectors. *Current protocols in human genetics* / editorial board, Jonathan L Haines [et al] *Chapter 12*, Unit 12 10.
- Sanders, R.W., Derking, R., Cupo, A., Julien, J.P., Yasmeen, A., de Val, N., Kim, H.J., Blattner, C., de la Pena, A.T., Korzun, J., *et al.* (2013). A next-generation cleaved, soluble HIV-1 Env trimer, BG505 SOSIP.664 gp140, expresses multiple epitopes for broadly neutralizing but not non-neutralizing antibodies. *PLoS pathogens* 9, e1003618.
- Sanders, R.W., van Gils, M.J., Derking, R., Sok, D., Ketas, T.J., Burger, J.A., Ozorowski, G., Cupo, A., Simonich, C., Goo, L., *et al.* (2015). HIV-1 neutralizing antibodies induced by native-like envelope trimers. *Science* 349, aac4223.
- Scharf, L., Wang, H., Gao, H., Chen, S., McDowall, A.W., and Bjorkman, P.J. (2015). Broadly Neutralizing Antibody 8ANC195 Recognizes Closed and Open States of HIV-1 Env. *Cell* 162, 1379-1390.
- Scheid, J.F., Mouquet, H., Feldhahn, N., Seaman, M.S., Velinzon, K., Pietzsch, J., Ott, R.G., Anthony, R.M., Zebroski, H., Hurley, A., *et al.* (2009). Broad diversity of neutralizing antibodies isolated from memory B cells in HIV-infected individuals. *Nature* 458, 636-640.
- Scheid, J.F., Mouquet, H., Ueberheide, B., Diskin, R., Klein, F., Oliveira, T.Y., Pietzsch, J., Fenyo, D., Abadir, A., Velinzon, K., *et al.* (2011). Sequence and structural convergence of broad and potent HIV antibodies that mimic CD4 binding. *Science* 333, 1633-1637.
- Shingai, M., Donau, O.K., Plishka, R.J., Buckler-White, A., Mascola, J.R., Nabel, G.J., Nason, M.C., Montefiori, D., Moldt, B., Poignard, P., *et al.* (2014). Passive transfer of modest titers of potent and broadly neutralizing anti-HIV monoclonal antibodies block SHIV infection in macaques. *The Journal of experimental medicine* 211, 2061-2074.
- Shingai, M., Nishimura, Y., Klein, F., Mouquet, H., Donau, O.K., Plishka, R., Buckler-White, A., Seaman, M., Piatak, M., Jr., Lifson, J.D., *et al.* (2013). Antibody-mediated immunotherapy of macaques chronically infected with SHIV suppresses viraemia. *Nature* 503, 277-280.
- Sok, D., Laserson, U., Laserson, J., Liu, Y., Vigneault, F., Julien, J.P., Briney, B., Ramos, A., Saye, K.F., Le, K., *et al.* (2013). The effects of somatic hypermutation on neutralization and binding in the PGT121 family of broadly neutralizing HIV antibodies. *PLoS pathogens* 9, e1003754.
- Walker, L.M., Huber, M., Doores, K.J., Falkowska, E., Pejchal, R., Julien, J.P., Wang, S.K., Ramos, A., Chan-Hui, P.Y., Moyle, M., *et al.* (2011). Broad neutralization coverage of HIV by multiple highly potent antibodies. *Nature* 477, 466-470.

- Walker, L.M., Phogat, S.K., Chan-Hui, P.Y., Wagner, D., Phung, P., Goss, J.L., Wrinn, T., Simek, M.D., Fling, S., Mitcham, J.L., *et al.* (2009). Broad and potent neutralizing antibodies from an African donor reveal a new HIV-1 vaccine target. *Science* 326, 285-289.
- West, A.P., Jr., Scharf, L., Scheid, J.F., Klein, F., Bjorkman, P.J., and Nussenzweig, M.C. (2014). Structural insights on the role of antibodies in HIV-1 vaccine and therapy. *Cell* 156, 633-648.
- Wu, X., Zhou, T., Zhu, J., Zhang, B., Georgiev, I., Wang, C., Chen, X., Longo, N.S., Louder, M., McKee, K., *et al.* (2011). Focused evolution of HIV-1 neutralizing antibodies revealed by structures and deep sequencing. *Science* 333, 1593-1602.
- Xiao, X., Chen, W., Feng, Y., Zhu, Z., Prabakaran, P., Wang, Y., Zhang, M.Y., Longo, N.S., and Dimitrov, D.S. (2009). Germline-like predecessors of broadly neutralizing antibodies lack measurable binding to HIV-1 envelope glycoproteins: implications for evasion of immune responses and design of vaccine immunogens. *Biochemical and biophysical research communications* 390, 404-409.
- Zhou, T., Georgiev, I., Wu, X., Yang, Z.Y., Dai, K., Finzi, A., Kwon, Y.D., Scheid, J.F., Shi, W., Xu, L., *et al.* (2010). Structural basis for broad and potent neutralization of HIV-1 by antibody VRC01. *Science* 329, 811-817.

Example 2: Supplemental Experimental Procedures to Example 1

[00459] DNA gene synthesis. Genes were synthesized at Genscript, Inc. Gp120 and gp140 variants in pHLsec contained a C-terminal GTKHHHHHH tag. Genes in pENTR contained a C-terminal cMyc epitope followed by a PDGFR transmembrane domain. IgGs were cloned into pFUSEss and Fabs were in a modified version of pFUSEss (pFABss). DNA was maxi-prepped using a BenchPro 2100.

[00460] Protein production. BG505-gp120 and variants based on BG505 contained the L111A mutation for more efficient production of monomer compared to other species (Hoffenberg et al., 2013) and the T332N mutation and were expressed in 293F cells grown in 293 Freestyle media (Life Technologies) by transient transfection with 293Fectin (Invitrogen). Protein was harvested from the supernatant 96 h post transfection and purified by nickel affinity chromatography on a HIS-TRAP column (GE) followed by HiLoad 16/600 Superdex 200 size exclusion chromatography (GE Healthcare). Gp140 SOSIPs were expressed in 293F cells grown in 293 Freestyle media by transient transfection with either 293Fectin or PEI. The protein was purified from the supernatant using a HIS-TRAP column, starting with a wash buffer (20 mM Imidazole, 500 mM NaCl, 20 mM Na₂HPO₄) and mixing with elution buffer (500 mM Imidazole, 500 mM NaCl, 20 mM Na₂HPO₄) using a linear gradient. The trimer fraction was collected and further purified on an S200Increase 10-

300 column (GE) in HBS (10 mM HEPES, 150 mM NaCl). The oligomeric state of the SOSIP trimers were then confirmed by size exclusion chromatography-multi-angle light scattering (SEC-MALS) using the DAWN HELEOS II multi-angle light scattering system with Optilab T-rEX refractometer (Wyatt Technology). The trimers were frozen in thin-walled PCR tubes at 1 mg/ml using liquid nitrogen and stored at -80°C (Jardine et al., 2015). Fabs and mAbs were produced in 293F cells as described previously (Jardine et al., 2013). For crystallography, SOSIP_MD39_10MUTA was expressed in 293S cells.

[00461] ELISA quantification of SOSIP expression. BG505 SOSIP variants were expressed using the Freestyle 293F expression system (Thermo Scientific) according to manufacturer's instructions. After 4 days, supernatants were harvested by centrifugation and stored at 4°C until analysis. Capture ELISAs were performed essentially as described previously (Schiffner et al., 2016). Briefly, ELISA plates were coated overnight with trimer specific PGT145 Fab at 4 µg/mL in PBS at 4°C followed by blocking with 2% w/v bovine serum albumin (BSA) in washing buffer (PBS + 0.05% v/v tween20). SOSIP expression supernatants were diluted 100x in sample buffer (washing buffer + 1% w/v BSA) and for each variant, a standard curve with known concentration of matching purified protein was prepared in sample buffer.

[00462] Supernatants and standard curves were added to ELISA plates and detected with trimer preferring IgG PGT151 at 10 µg/mL in sample buffer. Samples were labeled with horseradish peroxidase coupled Fcg-specific anti-human IgG (Jackson ImmunoResearch), developed and stopped with 1-Step Ultra TMB- ELISA substrate (Thermo Scientific) as per manufacturer's instructions, and optical densities were read at 450 nm and 570 nm. After background subtraction, data were fit to a "one-site specific binding with hill slope" curve in graphpad prism, and supernatant concentrations were extrapolated from standard curves.

[00463] Surface plasmon resonance (SPR). Kinetics and affinities of antibody-antigen interactions were measured on a ProteOn XPR36 (Bio-Rad) using GLC Sensor Chip (Bio-Rad) and 1x HBS-EP+ pH 7.4 running buffer (20x stock from Teknova, Cat. No H8022) supplemented with BSA at 1mg/ml. Human Antibody Capture Kit was used according to manufacturer's instructions (Cat. No BR-1008-39 from GE) to immobilize about 6000 RUs of capture mAb onto each flow cell. In a typical experiment, approximately 300-400 RUs of mAbs were captured onto each flow cell and analytes were passed over the flow cell at 50 µL/min for 3 min followed by a 5 min dissociation time. Regeneration was accomplished using 3M Magnesium Chloride with 180 seconds contact time and injected four times per

cycle. Raw sensograms were analyzed using ProteOn Manager software (Bio-Rad), including interspot and column double referencing, and either Equilibrium fits or Kinetic fits with Langmuir model, or both, were employed when applicable. Analyte concentrations were measured on a NanoDrop 2000c Spectrophotometer using Absorption signal at 280 nm (Jardine et al., 2015). Applicants measured kinetics and affinity of antibody-Fab-fragment antigen interactions on ProteOn XPR36 (Bio-Rad) using HTE Sensor Chip (Bio-Rad) and running buffer with 20 mM Sodium Phosphate Dibasic, pH 7.4, 500 mM Sodium Chloride, 50 mM Imidazole, supplemented with BSA at 1mg/ml and Tween 20 detergent at 0.05% v/v. Applicants used 0.1 M Nickel sulfate as activation solution. 0.5 M EDTA was Applicants' regeneration solution with 300 seconds contact time and injected two times per cycle (one time each for vertical and horizontal orientation). Raw sensograms were analyzed using ProteOn Manager software (Bio-Rad), interspot and column double referencing, Equilibrium or Kinetic with Langmuir model or both where applicable. Analyte concentrations were measured on NanoDrop 2000c Spectrophotometer using Absorption signal at 280 nm.

[00464] Design of PGT121 germline-targeting immunogens.

[00465] BG505-gp120 T332N fused to the PDGFR transmembrane domain (TM) was subjected to random mutagenesis using error prone PCR (gene morph II Agilent), and the resulting PCR product was gel purified and ligated into a modified version of the gateway cloning entry vector pENTR/D-TOPO (Ota et al., 2012) using the circular polymerase extension cloning (CPEC) method (Quan and Tian, 2014). The ligated vector containing the error prone library was purified using the PCR purification kit (Qiagen) and concentrated. The concentrated library was then transformed into electroMAX DH5a-E competent cells (Invitrogen) and grown overnight at 37°C in a 125 mL culture. The plasmid was purified using the BenchPro® 2100 (Invitrogen) and the gp120 insert was transferred to the lentiviral vector pLenti CMVTRE3G puro Dest (Ota et al., 2012) using the LR Clonase II enzyme mix (Invitrogen). The LR clonase reaction was scaled up ~10-fold to increase library size. The LR clonase product was again purified, concentrated and transformed into electroMAX stbl4 competent cells (Invitrogen) and grown overnight at 30°C in a 125 mL culture. This plasmid DNA was purified and ready for use in transfection. 293T cells cultured in Advanced DMEM (Gibco) supplemented with 5% FCS, GlutaMAX (Gibco), 2-mercaptoethanol (Gibco) and Antibiotic-Antimycotic (Gibco) were co-transfected with the BG505-gp120 error prone PCR library in pLenti CMVTRE3G puro Dest (10.8 µg), psPAX2 (7.0 µg) and pMD2.G (3.8 µg) with fugeneHD in a T75 flask (Salmon and Trono, 2007). The cells were kept at 37°C for

two days and then the media containing the virus was collected and spun down at 500g for 5min. 293T cells stably expressing rtTA3G from the pLenti CMV rtTA3G Blast vector (obtained from Dave Nemazee; (Ota et al., 2012)) were transduced at low moi (<0.1) in a T75 or T225 flask in the presence of 10 µg/mL blasticidin. The next day cells were selected with 2 µg/mL puromycin. 293T cells containing the stable library were induced with doxycycline (1µg/mL) and the following day were harvested in FACS buffer (HBSS, 1 mM EDTA, 0.5% BSA). Cells were stained with either the GL+9 or GL+3 Ab for ~30 min, washed with FACS buffer, and then stained with fluorescein isothiocyanate (FITC)-labeled α-cMyc (Immunology Consultants Laboratory) and phycoerythrin (PE)-conjugated α-human IgG (Sigma). Cells were sorted on a BD Influx (BD Biosciences) FACS sorter. Approximately 2×10^5 GL+9 positive cells were collected and expanded for ~one week in the presence of puromycin and blasticidin before the next round of enrichment was carried out. There was no enrichment for GL+3 positive cells after several rounds of sorting so only the GL+9 positive cells were sequenced. Once the desired population had been obtained the chromosomal DNA was extracted from the cell culture using the GenElute Mammalian Genomic DNA Miniprep Kit (Sigma). The BG505-gp120 gene was PCR amplified from the genomic DNA and ligated back into the Gateway entry vector using CPEC cloning and transformed into top10 competent cells. Later in the design process Gibson assembly was substituted for CPEC cloning. Colonies were sequenced at Genewiz. The sequences were highly enriched for two clones, one containing the N137 glycan knockout by the mutation T139I and the other containing the N133 glycan knockout by the mutation T135A in addition to the T139I mutation. These constructs were called 2MUT (T332N, T139I) and 3MUT (T332N, T135A, T139I). Measuring the affinities of gp120-2MUT and gp120-3MUT against a panel of partially mutated PGT121 Abs (table S1) indicated that knocking out both glycans gave a larger boost in affinity compared to only the N137 glycan-KO so 3MUT was used for further designs.

[00466] In parallel to screening the error prone PCR library, a combinatorial library was created based on the structure of PGT122 in complex with BG505 SOSIP (PDB IDs 4NCO and 3J5M). Because the initial SOSIP structures were low resolution and structures of germline PGT121 showed light chain conformational changes Applicants elected to do a saturation mutagenesis combinatorial library that would roughly cover the length of the V1 loop that could potentially interact with germline PGT121 Abs. The library was generated by PCR amplifying the BG505 SOSIP construct in two partially overlapping fragments. The C-

terminal fragment was amplified with a primer containing the degenerate codon NNK at four positions in the V1 loop (V134, N136, I138, and D140) as well as a degenerate base encoding N or D at position 137. The two PCR products were ligated together using a second round of PCR, and this second PCR product was inserted into the pENTR vector as described above. The resulting construct was transferred to the pLenti CMVTRE3G puro Dest vector, and lentivirus was produced. Stable cells were stained with the GL+3 Ab and α -cMyc, and double positive cells were sorted. This resulted in a binding population that was sequenced and found to be a single unique clone containing the mutations V134Y, N136P, I138L, D140N. This clone was called 5MUT (T332N, V134Y, N136P, I138L, D140N). These mutations were combined with the T139I mutation (6MUT) or the T135A/T139I mutations (7MUT).

[00467] Next, a saturation mutagenesis scanning library was created on the gp120-7MUT construct using site directed mutagenesis with the QuikChange kit (Agilent Technologies) with a unique NNK/MNN primer pair for each position that was scanned. 11 positions in the V1 loop (T132 to M142) and 10 positions in the V3 loop (T320 to Q328) were scanned and the resulting 21 reactions were pooled, purified, concentrated, and transformed into electroMAX DH5a-E competent cells and transferred to pLenti CMVTRE3G puro Dest as described above. This library was then stained separately with GL_{CDR3mat}, GL_{CDR3rev4}, or a Chimeric Ab containing the mature PGT121 heavy chain paired with the GL_{CDR3rev4} light chain (121H/GLL-rev4), as well as α -cMyc for expression. Double positive cells were sorted and 3 mutations were enriched in the GL_{CDR3mat} sort (N137F, T320F, Q328M) and two mutations were enriched in the 121H/GLL-rev4 sort (N135R, Q328M) whereas a binding population was not obtained in the GL_{CDR3rev4} sort. Combining these mutations with 7MUT resulted in 9MUT_A (7MUT + N137F/Q328M), 9MUT_B (7MUT + N135R/Q328M), and 10MUT (7MUT + N137F/T320F/Q328M). The 9MUT_B protein showed improved binding to 121H/GLL-rev4 but worse binding to all other PGT121-class antibodies tested compared to 7MUT (from which 9MUT_B was derived) and so 9MUT_B was not selected for further use except as a control for the chimeric antibody (data not shown). Gp120-10MUT showed better binding to GL_{CDR3mat} compared to gp120-9MUT_A and T320F was used in subsequent designs with the exception of Applicants' SOSIP-10MUT crystal structure, which lacks the T320F mutation.

[00468] Having established ~1 μ M binding to the GL_{CDR3mat} Ab with 10MUT Applicants' goal was to improve the immunogen to tolerate more variation within the H-CDR3. For this Applicants created three more V1 loop combinatorial libraries each containing four NNK

codons. The three libraries contained NNK codons at positions (A135/P136/F137/L138), (F137/L138/I139/N140), and (I139/N140/D141/M150). Each library was assembled from two partially overlapping ultramers (Integrated DNA Technologies) and ligated into the gp120-10MUT gene using gibbon assembly (New England Biolabs). The three libraries were pooled and screened against zup GL_{CDR3rev2} and GL_{CDR3rev4} Abs. Sorting against the GL_{CDR3rev4} Ab resulted in enrichment for the D141N mutation (11MUT_A) and sorting against the GL_{CDR3rev2} resulted in enrichment for L139 and S140 with the most frequent clone containing the sequence N137/L138/L139/S140. When these mutations were combined with the D141N mutation as well as a T415V mutation, which Applicants had identified as being beneficial for binding to PGT121 on an engineered outer domain construct (data not shown), it resulted in 11MUT_B.

Development of BG505-SOSIP_{MD39}.

[00469] BG505 SOSIP “rare amino acid” library. The BG505 SOSIP “rare amino acid” library was synthesized at GenScript. It was first sorted against PG16 followed by a sort for a high PGT145/B6 binding ratio. The cells were expanded for 1 week and then sorted for either high PGT145/B6 or high PGT151/4025. After six rounds of sorting the library was sequenced (Genewiz). PGT145, PGT151, and PG16 Fabs contained HA epitope tags and were labeled with α -HA-PE (Miltenyi Biotec). B6 and 4025 Fabs contained V5 epitope tags and were labeled with α -V5-FITC (GeneTex).

[00470] BG505 SOSIP whole gene saturation mutagenesis. The whole gene saturation mutagenesis library was synthesized at Integrated DNA Technologies in four segments that each contained ~150 NNK codons that were cloned into the BG505-SOSIP gene using either CPEC or Gibson assembly which resulted in four libraries. NNK codons were barcoded with a silent mutation on each side. The libraries created from the second and third segments were combined into one. The first, second and third libraries had NNK codons covering residues Y39-N185, N186-R500 and K502-Q658, respectively. The library that covered gp41 (502-658) was sorted for high PGT145/cMyc, high PGT145/B6, and high PGT151/cMyc. The first gp120 library (39-185) was sorted for high PGT145/B6, and high PGT151/4025. The second gp120 library (186-

[00471] 500) was sorted for high PGT145/cMyc, high PGT145/B6, high PGT151/4025, and high PGT151/cMyc. The sorted libraries were sequenced and analyzed essentially as described previously (Jardine et al., 2016). Positions that enriched for the same amino acid against multiple different mAb sorts (E.g. PGT145(+)/B6(-) and PGT151(+)/4025(-)) were

avored for testing in follow up combinatorial libraries or directly testing in recombinantly purified protein. Combinatorial libraries based on the next generation sequencing analysis were assembled from overlapping ultramers and sorted against the same antibodies described above.

[00472] Differential scanning calorimetry (DSC). DSC experiments were performed on a MicroCal VP-Capillary differential scanning calorimeter (Malvern Instruments). The HEPES buffered saline (HBS) buffer was used for baseline scans and the protein samples were diluted into HBS buffer to adjust to 0.25 mg/ml. The system was allowed to equilibrate at 20 °C for 15 min and then heat up till 90°C at a scan rate of 90°C/h. Buffer correction, normalization, and baseline subtraction were applied during data analysis using Origin 7.0 software. The non-two-state model was used for data fitting.

[00473] ELISA to characterize antigenic profile of native-like trimers. 96-well plates were coated overnight at 4°C with 6x-His Epitope Tag Antibody (Thermofisher) at 2 mg/ml in PBS. Plates were washed 3 times with PBS, 0.05% Tween (PBS-T), and blocked with 10% milk PBS for 1h. Subsequently, 2 mg/ml of the purified His-tagged SOSIP protein was added for 2 h in 1% milk PBS-T, after which the plates were washed three times with PBS-T. Serial dilutions of mAbs in 1% milk PBS-T were added to the plates for 1 h, after which the plates were washed again three times with PBS-T before the addition of anti-human Fc region-conjugated alkaline phosphatase (Jackson ImmunoResearch) at 1:1000 for 1 h. After four final washes, binding was detected by the addition of alkaline phosphatase substrate and measured by absorbance at 405 nm.

[00474] Development of variable loop cocktail (VLC) trimers. Using BG505 SOSIP MD39 trimer as a base, a series of new trimers were engineered by replacing the immunodominant variable loops of the BG505 strain with loops from alternative strains. Given the vast number of HIV strains available, Applicants created three separate criteria to guide loop selection. For the first set of variable loop transplants, Applicants cataloged the number of glycans within each variable loop and the length of each variable loop (Figure 7C). Certain combinations of variable loop lengths and glycans were observed more frequently than others across HIV strains (e.g. 20.48 % of HIV strains have a 14 amino acid variable loop 2 with one glycan, Figure 7C). Applicants searched for strains that contain the most common loop length/glycan combination for each of the variable loops (V1,V2,V4,V5). For the second set of variable loop transplants, Applicants searched for strains with variable loops of the same length and number of glycans as BG505, but with very different amino acid

sequence and glycan positioning within the loops as compared to BG505. No single strain had all the same variable loop lengths and number of glycans as BG505, so Applicants relaxed criteria and matched each variable loop independently for this set of variable loop transplants only. For the third set of variable loop transplants, Applicants searched for strains with exceptionally long variable loops (V1,V2,V4 must be ≥ 4 amino acids longer than BG505). Under each of these criteria, Applicants were able to obtain one or two trimers that had a reasonable level of expression and formed well-behaved native-like trimers (Figure 14A). The loops of the VLCs are defined as: VLC-1 (V1: BES10.EF363127, V2: BL8157.DQ886035, V4: BF1P51.JQ250880, V5: CZM197.DQ388515), VLC-2(BG505.DQ208458), VLC-3 (PRLS08.FJ469757) VLC-4(GHJ193.AB231897), VLC-5(OUR2478P.EF165541). A region defined as 335-351 (HxB2) underneath V4 was included when transplanting V4, due to high variability and close contact with V4. Including BG505, Applicants report a set of 5 trimers with diverse variable loops. One version of the VLC trimers that did not have the MD39 mutations, and instead contained an extra stabilizing disulfide (DS21: V120C-Q315C) in order to staple down the tip of the V3, this version of the VLC trimers was used in an accompanying manuscript (Escolano et al., 2016).

[00475] Supplemental References

- Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., *et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D* 66, 213-221.
- Hoffenberg, S., Powell, R., Carpov, A., Wagner, D., Wilson, A., Kosakovsky Pond, S., Lindsay, R., Arendt, H., Destefano, J., Phogat, S., *et al.* (2013). Identification of an HIV-1 clade A envelope that exhibits broad antigenicity and neutralization sensitivity and elicits antibodies targeting three distinct epitopes. *J Virol* 87, 5372-5383.
- Kulp, D.W., Subramaniam, S., Donald, J.E., Hannigan, B.T., Mueller, B.K., Grigoryan, G., and Senes, A. (2012). Structural informatics, modeling, and design with an open-source Molecular Software Library (MSL). *Journal of computational chemistry* 33, 1645-1661.
- McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. *J Appl Crystallogr* 40, 658-674.
- Murshudov, G.N., Skubak, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn, M.D., Long, F., and Vagin, A.A. (2011). REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D* 67, 355-367.

- Ogura, T., Iwasaki, K., and Sato, C. (2003). Topology representing network enables highly accurate classification of protein images taken by cryo electron-microscope without masking. *Journal of structural biology* 143, 185-200.
- Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Method Enzymol* 276, 307-326.
- Quan, J., and Tian, J. (2014). Circular polymerase extension cloning. *Methods in molecular biology* 1116, 103-117.
- Ringe, R.P., Sanders, R.W., Yasmeen, A., Kim, H.J., Lee, J.H., Cupo, A., Korzun, J., Derking, R., van Montfort, T., Julien, J.P., *et al.* (2013). Cleavage strongly influences whether soluble HIV-1 envelope glycoprotein trimers adopt a native-like conformation. *Proceedings of the National Academy of Sciences of the United States of America* 110, 18256-18261.
- Salmon, P., and Trono, D. (2007). Production and titration of lentiviral vectors. *Current protocols in human genetics* / editorial board, Jonathan L Haines [et al] *Chapter 12*, Unit 12 10.
- Schiffner, T., de Val, N., Russell, R.A., de Taeye, S.W., de la Pena, A.T., Ozorowski, G., Kim, H.J., Nieusma, T., Brod, F., Cupo, A., *et al.* (2016). Chemical Cross-Linking Stabilizes Native-Like HIV-1 Envelope Glycoprotein Trimer Antigens. *Journal of virology* 90, 813-828.

[00476]

[00477]

* * *

[00478] Having thus described in detail preferred embodiments of the present invention, it is to be understood that the invention defined by the above paragraphs is not to be limited to particular details set forth in the above description as many apparent variations thereof are possible without departing from the spirit or scope of the present invention.

[00479] Although the present invention and its advantages have been described in detail, it should be understood that various changes, substitutions and alterations can be made herein without departing from the spirit and scope of the invention as defined in the appended claims.

WHAT IS CLAIMED IS:

1. A non-naturally occurring protein comprising an immunogen comprising a gp120 or gp140 trimer comprising one or more stabilizing mutations, cleavage-independent modifications, and/or an anchored membrane.
2. The non-naturally occurring protein of claim 1 comprising:
 - (a) a Type I or Type II immunogen, wherein the Type I immunogen is a gp120 and the Type II immunogen is a gp140 trimer molecules with one or more mutations that improve binding to germline-reverted and/or less-mutated versions of PGT121 or
 - (b) a Type III immunogen, wherein the Type III immunogen is a gp140 trimer molecule with one or more stabilizing mutations to increase expression level and/or increase thermal melting temperature and/or improve antigenic profile, where a favorable antigenic profile means better affinity for broadly neutralizing antibodies and no or very weak affinity for non-neutralizing antibodies or
 - (c) a Type IV immunogen, wherein the Type IV immunogen is a combination of one or more mutations from the Type II immunogen and Type III immunogen, wherein the Type IV immunogen is a gp140 trimer comprising both stabilizing mutations and germline-targeting mutations or
 - (d) a Type V immunogen, wherein the Type V immunogen is a trimer with a modified surface or of a strain other than BG505 or
 - (e) a Type VI immunogen, wherein the Type VI immunogen comprises one or more additional trimer modifications that add extra functionality and that can be combined with Types II, III, IV or V or
 - (f) a Type VII immunogen, wherein the Type VII immunogen is a native-like trimer from other HIV strains stabilized by MD39 and Olio6 mutations or
 - (h) a Type VIII immunogen, wherein the Type VIII immunogen is a cleavage-independent trimer which is a variant of BG505 MD39 that does not require cleavage by furin or
 - (i) a Type IX immunogen, wherein the Type IX immunogen is a glycan masked trimer in which N-linked glycosylation sites cover the bottom and sides of the soluble trimer or
 - (j) a Type X immunogen, wherein the Type X immunogen is a native-like trimer with variable loops V1, V2b and V4 modified to minimize their lengths and maximize the number of glycosylation sites contained within them or

(i) a Type XI immunogen, wherein the Type XI immunogen is a BG505 MD39-based, single-component, self-assembling nanoparticle or

(j) a Type XII immunogen, wherein the Type XII immunogen is a BG505 MD39-based, membrane-bound native like trimer.

3. The non-naturally occurring protein of claim 1 or 2, wherein the protein comprises any one of:

(a) BG505-gp120-L111A-2_T135A_T139I_mC (BG505-gp120 3mut)

VWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQEIHLNVTEEFNMWKNMVEQMH
TDIISAWDQSLKPCVKLTPLCVTLQCTNVANNIIDDMRGELKNCSFNMTTEL RDKKQKVYSL
FYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCK
DKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEEVMIRSENITNNAKNILVQFNTF
VQINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNVSKATWNETLGKVVKQLRKHFG
NNTIIRFANSSGGDLEVTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQGSNSTGSNDSITLP
CRIKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGGGDMRDNW
RSELYKVKVKIEP

(b) BG505-gp120-L111A-2_5mut_mC (BG505-gp120 5mut)

VWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQEIHLNVTEEFNMWKNMVEQMH
TDIISAWDQSLKPCVKLTPLCVTLQCTNYTPNLTNDMRGELKNCSFNMTTEL RDKKQKVYSL
FYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCK
DKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEEVMIRSENITNNAKNILVQFNTF
VQINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNVSKATWNETLGKVVKQLRKHFG
NNTIIRFANSSGGDLEVTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQGSNSTGSNDSITLP
CRIKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGGGDMRDNW
RSELYKVKVKIEP

(c) BG505-gp120-L111A-2_7mut_mC (BG505-gp120 7mut or BG505 gp120 7MUT)

VWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQEIHLNVTEEFNMWKNMVEQMH
TDIISAWDQSLKPCVKLTPLCVTLQCTNYAPNLIINDMRGELKNCSFNMTTEL RDKKQKVYSL
FYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCK
DKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEEVMIRSENITNNAKNILVQFNTF
VQINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNVSKATWNETLGKVVKQLRKHFG
NNTIIRFANSSGGDLEVTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQGSNSTGSNDSITLP
CRIKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGGGDMRDNW
RSELYKVKVKIEP

(d) BG505-gp120-L111A-2_10mut_mC (BG505-gp120 10mut or BG505 gp120 10MUT)

VWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQEIHLNVTEEFNMWKNMVEQMH
TDIISAWDQSLKPCVKLTPLCVTLQCTNYAPFLIINDMRGELKNCSFNMTTEL RDKKQKVYSL
FYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCK

DKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEEVVMIRSENITNNAKNILVQFNTF
VQINCTRPNNNTRKSIRIGPGQAFYA**FGDIIGDIRMA**HNCNVSKATWNETLGKVVQQLRKHF
GNTTIIRFANSSGGDLEVTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQGSNSTGSNDSITLP
CRIKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGGGDMRDNW
RSELYKYKVVKIEP

(e) BG505-gp120-L111A-2_10mut2A_mC

VWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQEIHLNVTEEFNMWKNMVEQMH
TDIISAWDQSLKPCVKLTPLCVTLQCTNY**APFLIN**MRGELKNCSFNMTTELDDKKQKVYSL
FYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCK
DKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEEVVMIRSENITNNAKNILVQFNTF
VQINCTRPNNNTRKSIRIGPGQAFYA**FGDIIGDIRMA**HNCNVSKATWNETLGKVVQQLRKHF
GNTTIIRFANSSGGDLEVTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQGSNSTGSNDSITLP
CRIKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGGGDMRDNW
RSELYKYKVVKIEP

(f) BG505-gp120-L111A-2_11mut2A_mC

VWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQEIHLNVTEEFNMWKNMVEQMH
TDIISAWDQSLKPCVKLTPLCVTLQCTNY**APNLLSN**MRGELKNCSFNMTTELDDKKQKVYSL
FYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCK
DKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEEVVMIRSENITNNAKNILVQFNTF
VQINCTRPNNNTRKSIRIGPGQAFYA**FGDIIGDIRMA**HNCNVSKATWNETLGKVVQQLRKHF
GNTTIIRFANSSGGDLEVTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQGSNSTGSNDSI**VLP**
CRIKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGGGDMRDNW
RSELYKYKVVKIEP

4. The non-naturally occurring protein of claim 1 or 2r, wherein the trimer comprises any one of:

(a) BG505_SOSIP.D664_JS_3mut_mC (SOSIP-3MUT)

AENLWVTVYYGVVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQEIHLNVTEE
FNMWKNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLQCTNV**ANNI**IDDMRGELKNCSFNMT
TELDDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIH
YCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEEVVMIRSENIT
NNAKNILVQFNTFVQINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHNCNVSKATWNET
LGKVVQQLRKHF
GNTTIIRFANSSGGDLEVTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQGSNSTGSNDSITLP
CRIKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGGGDMRDNW
RSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRAVGIGAVFLGFLGAA
GSTMGAASMTLTQARNLLSGIVQQQSNLLRAPEAQQHLLKLTWVGKQLQARVLAVERYLR
DQQLLGIWGC SGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQ
NQQEKNEQDLLALD

(b) BG505_SOSIP.D664_JS_5mut_mC (SOSIP-5MUT)

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLNVTEE
 FNMWKNNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLQCTN**YTPNLTND**MRGELKNCSFNMT
 TELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIH
 YCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEEVVMIRSENIT
 NNAKNILVQFNTVPVQINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHNCNVSKATWNET
 LGKVVKQLRKHFNGNTIIRFANSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQG
 SNSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQGVIKRVSNITGLILTRDGGSTNSTTE
 TFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRAVGIGAVFLGFLGAA
 GSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQQHLLKLTWVGIKQLQARVLAVERYLR
 DQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQ
 NQQEKNEQDLLALD

(c) BG505_SOSIP.D664_JS_7mut_mC (SOSIP-7MUT)

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLNVTEE
 FNMWKNNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLQCTN**YAPNLIND**MRGELKNCSFNMT
 TELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIH
 YCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEEVVMIRSENIT
 NNAKNILVQFNTVPVQINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHNCNVSKATWNET
 LGKVVKQLRKHFNGNTIIRFANSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQG
 SNSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQGVIKRVSNITGLILTRDGGSTNSTTE
 TFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRAVGIGAVFLGFLGAA
 GSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQQHLLKLTWVGIKQLQARVLAVERYLR
 DQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQ
 NQQEKNEQDLLALD

(d) BG505_SOSIP.D664_JS_10mut_mC (SOSIP-10MUT)

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLNVTEE
 FNMWKNNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLQCTN**YAPFLIND**MRGELKNCSFNMT
 TELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIH
 YCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEEVVMIRSENIT
 NNAKNILVQFNTVPVQINCTRPNNNTRKSIRIGPGQAFYA**FGDIIGDIRMA**HNCNVSKATWNET
 LGKVVKQLRKHFNGNTIIRFANSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQG
 SNSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQGVIKRVSNITGLILTRDGGSTNSTTE
 TFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRAVGIGAVFLGFLGAA
 GSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQQHLLKLTWVGIKQLQARVLAVERYLR
 DQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQ
 NQQEKNEQDLLALD

(e) BG505_SOSIP.D664_JS_10mut2A_mC

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLENVTEE
 FNMWKNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLQCTNY**APFLIN**NMRGELKNCSFNMT
 TELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPPIPIH
 YCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLNGSLAEEVVMIRSENIT
 NNAKNILVQFNTVPVQINCTRPNNNTRKSIRIGPGQAFYA**FGDI**IGDIR**MAH**CNVSKATWNET
 LGKVVKQLRKHFGNNTIIRFANSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQG
 SNSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQGVIKRVSNITGLILTRDGGSTNSTTE
 TFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVFLGFLGAA
 GSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQQHLLKLTWGIKQLQARVLAVEHYLR
 DQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQ
 NQQEKNEQDLLALD

(f) BG505_SOSIP.D664_JS_11mut2A_mC

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLENVTEE
 FNMWKNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLQCTNY**APNLLS**NMRGELKNCSFNMT
 TELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPPIPIH
 YCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLNGSLAEEVVMIRSENIT
 NNAKNILVQFNTVPVQINCTRPNNNTRKSIRIGPGQAFYA**FGDI**IGDIR**MAH**CNVSKATWNET
 LGKVVKQLRKHFGNNTIIRFANSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQG
 SNSTGSNDSI**VL**PCRIKQIINMWQRIGQAMYAPPIQGVIKRVSNITGLILTRDGGSTNSTTE
 TFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVFLGFLGAA
 GSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQQHLLKLTWGIKQLQARVLAVEHYLR
 DQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQ
 NQQEKNEQDLLALD

5. The non-naturally occurring protein of claim 1 or 2, wherein the trimer comprises any one of:

(a) BG505_SOSIP_D664_MD39_mC (BG505 SOSIP-MD39 or MD39)

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLENVTEE
 FNMWKNMVEQM**H**EDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELKNCSFNMT
 TELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPPIPIH
 YCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLNGSLAEEV**I**IRSENIT
 NNAKNILVQ**L**NTPVQINCTRPNNNT**V**KSIRIGPGQAF**Y**YT**GDI**IGDIR**QA**HCNVSKATWNET
 LGKVVKQLRKHFGNNTIIR**FA**QSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQG
 SNSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQGVIKRVSNITGLILTRDGGSTNSTTE
 TFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAV**SL**GFLGAA
 GSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPE**P**QQHLL**KD**THWGIKQLQARVLAVE**H**YLR
 DQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQ
 NQQEKNEQDLLALD

(b) BG505_SOSIP_D664_MD16_mC stabilizes the V3 loop

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLENVTEE
 FNMWKNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELKNCSFNMT
 TELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPPIPIH
 YCAPAGWAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLNGLSLAESEVMIRSENIT
 NNAKNILVQFNTFPVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNVSKATWNET
 LGKVVKQLRKHFNGNTIIRFANSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQG
 SNSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTE
 TFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVFLGFLGAA
 GSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQQHLLKDTVWGIKQLQARVLAVEHYLR
 DQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQ
 NQQEKNEQDLLALD

(c) BG505_SOSIP_D664_split1_1_mC increases melting temp 10C

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLENVTEE
 FNMWKNMVEQMHTDIIELWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELKNCSFNMT
 TELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPPIPIH
 YCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLNGLSLAESEVMIRSENIT
 NNAKNILVQFNTFPVQINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNVSKATWNET
 LGKVVKQLRKHFNGNTIIRFANSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQG
 SNSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTE
 TFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVFLGFLGAA
 GSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQQHLLKLTWGIKQLQARVLAVEHYLR
 DQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQ
 NQQEKNEQDLLALD

(d) BG505_SOSIP_D664_MD39C_mC highest melting temp 82.5C

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLENVTEE
 FNMWKNMVEQMHTDIIELWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELKNCSFNMT
 TELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPPIPIH
 YCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLNGLSLAESEVIIRSENIT
 NNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNVSKATWNET
 LGKVVKQLRKHFNGNTIIRFAQSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQG
 SNSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTE
 TFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVSLGFLGAA
 GSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKLTWGIKQLQARVLAVEHYLR
 DQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQ
 NQQEKNEQDLLALD

(e) BG505_SOSIP_D664_MD9_mC improved yield/stability

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLENVTEE
 FNMWKNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELKNCSFNMT

TEL RDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIH
 YCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLNGSLAEFEVMIRSENIT
 NNAKNILVQFNTVPVQINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHNCVSKATWNET
 LGKVVKQLRKHFNGNTIIRFANSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQG
 SNSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQGIVIRCVSNITGLILTRDGGSTNSTTE
 TFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVFLGFLGAA
 GSTMGAASMTLTQARNLLSGIVQQQSNLLRAPE**R**QQHLLKDTVWGIKQLQARVLAVE**H**YLR
 DQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQ
 NQQEKNEQDLLALD

(f) BG505_SOSIP_D664_MD2_mC improved yield

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLNVTEE
 FNMWKNMVEQMHTDIIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELKNCSFNMT
 TEL RDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIH
 YCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLNGSLAEFEVMIRSENIT
 NNAKNILVQFNTVPVQINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHNCVSKATWNET
 LGKVVKQLRKHFNGNTIIRFANSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQG
 SNSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQGIVIRCVSNITGLILTRDGGSTNSTTE
 TFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVFLGFLGAA
 GSTMGAASMTLTQARNLLSGIVQQQSNLLRAPE**A**QQHLLKDTVWGIKQLQARVLAVE**R**YLR
 DQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQ
 NQQEKNEQDLLALD

(g) BG505_SOSIP_D664_olio6_mC

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLNVTEE
 FNMWKNMVEQMHEIDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELKNCSFNMT
 TEL RDKKQKVYSLF**W**RLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIH
 YCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLNGSLAEFEV**I**IRSENIT
 NNAKNILVQLNTPVQINCT**A**PN**N**F**T**VKSIRIGPGQAFY**M**GDIIGDIRQAHNCVSKATWNET
 LGKVVKQLRKHFNGNTIIR**F**AQSSGGDLEVTTTHSFNCG**G**MFFYCNTSGLFNSTWISNTSVQG
 SNSTGSNDSITLPCRI**K**LIINMWQRIGQAMYAPPIQGIVIRCVSNITGLILTRDGGSTNSTTE
 TFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVSLGFLGAA
 GSTMGAASMTLTQARNLLSGIVQQQSNLLRAPE**P**QQHLLKDTHWGIKQLQARVLAVE**H**YLR
 DQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQ
 NQQEKNEQDLLALD

(h) BG505_SOSIP_D664_MD53_mC

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLNVTEE
 FNMWKNMVEQMHEIDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELKNCSFNMT
 TEL RDKKQKVYSLF**W**RLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIH
 YCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKP**V**LSTQLLNGSLAEFEV**I**VRSENIT

NNAKNILVQLNTPVQINCTAPNNFTVKSIRIGPGQAFYYMGDIIGDIRQAHNCNVSKATWNET
 LGKVVKQLRKHFNGNTIIRFAQSSGGDLEVTTTHSFNCGGMFFFCNTSGLFNSTWISNTSVQG
 SNSTGSNDSITLPCRIKLIINMWQRIGQAMYAPPIQGIVRCVSNITGLILTRDGGSTNSTTE
 TFRPGGGDMRDVWRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVSLGFLGAA
 GSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARVLAVEHYLR
 DQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQ
 NQQEKNEQDLLALD

(i) BG505_SOSIP_D664_MD37

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLENVTEE
 FNMWKNMVEQMHEIDIISLWDQSLKPCCKLTPLCVTLQCTNVTNNITDDMRGELKNCSFNMT
 TELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSACTQACPKVSFEPPIPIH
 YCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGLSLAEEVEIIRSENIT
 NNAKNILVQLNTPVQINCTRPNNNTRKSIRIGPGCAFYATGDIIGDIRQAHNCNVSKATWNET
 LGKVVKQLRKHFNGNTIIRFAQSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQG
 SNSTGSNDSITLPCRIKQIINMWQRIGQCMYAPPIQGIVRCVSNITGLILTRDGGSTNSTTE
 TFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVSLGFLGAA
 GSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQEHLLKDTHWGIKQLQARVLAVEHYLR
 DQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQ
 NQQEKNEQDLLALD

6. A The non-naturally occurring protein of claim 1 or 2, wherein the trimer comprises any one of:

(a) BG505_SOSIP_D664_MD39_9mut2A_mC (MD39+9MUT)

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLENVTEE
 FNMWKNMVEQMHEIDIISLWDQSLKPCVKLTPLCVTLQCTNYAPFLINNMARGELKNCSFNMT
 TELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPPIPIH
 YCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGLSLAEEVEIIRSENIT
 NNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRMAHNCNVSKATWNET
 LGKVVKQLRKHFNGNTIIRFAQSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQG
 SNSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQGIVRCVSNITGLILTRDGGSTNSTTE
 TFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVSLGFLGAA
 GSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARVLAVEHYLR
 DQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQ
 NQQEKNEQDLLALD

(b) BG505_SOSIP_D664_MD39_10mut_mC

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLENVTEE
 FNMWKNMVEQMHEIDIISLWDQSLKPCVKLTPLCVTLQCTNYAPFLINDMRGELKNCSFNMT
 TELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPPIPIH
 YCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGLSLAEEVEIIRSENIT

NNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPGQAFYYFGDIIGDIRMAHCNVSKATWNET
 LGKVVKQLRKHFNGNTIIRFAQSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQG
 SNSTGSNDSITLPCRICKIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTE
 TFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVSLGFLGAA
 GSTMGAASMTLTQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARVLAVEHYLR
 DQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQ
 NQQEKNEQDLLALD

(c) BG505_SOSIP_D664_MD39_10mut2A_mC

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQEIHLENVTEE
 FNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNYAPFLINNMARGELKNCSFNMT
 TELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPPIPIH
 YCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLNGSLAEEVEIIRSENIT
 NNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPGQAFYYFGDIIGDIRMAHCNVSKATWNET
 LGKVVKQLRKHFNGNTIIRFAQSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQG
 SNSTGSNDSITLPCRICKIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTE
 TFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVSLGFLGAA
 GSTMGAASMTLTQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARVLAVEHYLR
 DQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQ
 NQQEKNEQDLLALD

(d) BG505_SOSIP_D664_MD39_11mut2A_mC

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQEIHLENVTEE
 FNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNYAPNLLSNMARGELKNCSFNMT
 TELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPPIPIH
 YCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLNGSLAEEVEIIRSENIT
 NNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPGQAFYYFGDIIGDIRMAHCNVSKATWNET
 LGKVVKQLRKHFNGNTIIRFAQSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQG
 SNSTGSNDSIVLPCRICKIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTE
 TFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVSLGFLGAA
 GSTMGAASMTLTQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARVLAVEHYLR
 DQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQ
 NQQEKNEQDLLALD

7. The non-naturally occurring protein of claim 1 or 2, that can be employed in strategic boosting regimens, wherein the trimer comprises any one of:

(a) BG505_SOSIP_MD39_VLC1-03

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQEIHLENVTEE
 FNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTDWDNATLANMTGEIKNCSFNMT
 TELRDKKQKVYSLFYELDIIP IENEYISNNNTSNTSYRLINCNTSAITQACPKVSFEPPIPIH
 YCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLNGSLAEEVEIIRSENIT

NNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHNCV**SATQWEQT**
LKGIAAKLLEHFGNNTIIRFAQSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNG**SSWNLNKTKE**
NTTNLENGTITLPCRIKQIINMWQRIGQAMYAPPIQG**VIRCVSNITGLILTRDGGNKSAGIE**
 TFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVSLGFLGAA
 GSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARVLAVEHYLR
 DQQLLGIWGC**SGKLICCTNVPWNSSWSNRNLSEI**WDNMTWLQWDKEISNYTQIIYGLLEESQ
 NQQEKNEQDLLALD

(b) BG505_SOSIP_MD39_VLC2-04

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLNVTEE
 FNMWKNMVEQM**HEDIISLWDQSLKPCVKLTPLCVTLQCS**DYEGNT**TRQNITMKEEKGEIKN**
 CSFNMTTEL**RDKKQKVYSLFYKLDITPIEEDNNSNNSSSANS**SSNSNANYTNYRLINCNTSAI
 TQACPKVSFEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQ**LLNG**
 SLAEEV**II**RSENITNNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDI
 RQAHNCV**SGTKWKNTLQ**IVK**KLGDH**FGNNTIIRFAQSSGGDLEVTTTHSFNCGGEFFYCNTS
 GLFNSTW**NRTNGTWNDVEGLNYTNGNDT**ITLPCRIKQIINMWQRIGQAMYAPPIQG**VIRCVS**
 NITGLILTRDGG**NDTDK**NETFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRR
 RRRRAVGIGAVSLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDT
 HWGIKQLQARVLAVEHYLRDQQLLGIWGC**SGKLICCTNVPWNSSWSNRNLSEI**WDNMTWLQW
 DKEISNYTQIIYGLLEESQNQQEKNEQDLLALD

(c) BG505_SOSIP_MD39_VLC2-08

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLNVTEE
 FNMWKNMVEQM**HEDIISLWDQSLKPCVKLTPLCVTLQCVTLKNC**SNSNC**SISRNIS**IE**MDG**
EIKNCSFNMTTEL**RDKKQKVYSLFYRLDIVPIESSNNSQLSNNSQVSNN**SQSSNYSQYRLIN
 CNTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVST
 QLLNGSLAEEV**II**RSENITNNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPGQAFYYTG
 DIIGDIRQAHNCV**SKDWEKTLQQVATKLGQH**FGNNTIIRFAQSSGGDLEVTTTHSFNCGGEF
 FYCNTSGLFNSTW**IRSSNSTWNSSASNSTELNSNITLPCRIKQIINMWQRIGQAMYAPPIQ**
 GVIRCVSNITGLILTRDGG**HETENKT**ETFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRC
 KRRVVGRRRRRRRAVGIGAVSLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEP
 QQHLLKDTHWGIKQLQARVLAVEHYLRDQQLLGIWGC**SGKLICCTNVPWNSSWSNRNLSEI**W
 DNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALD

(d) BG505_SOSIP_MD39_VLC3-13

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLNVTEE
 FNMWKNMVEQM**HEDIISLWDQSLKPCVKLTPLCVTLQCTNAALT**NVTITNGPNITEEIRNC
 SFNMTTEL**RDKKQKVYSLFYKLDLVQINGSGGEYRLINCNTSAITQACPKVSFEPIPIHYCA**
 PAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLNGSLAEEV**II**RSENITNNA
 KNILVQLNTPVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHNCV**SGTKWNETLKQ**
VAGKLRKHFGNNTIIRFAQSSGGDLEVTTTHSFNCGGEFFYCNTSGLFN**STWPENG**TMEGSNG

TITLPCRICKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGG**NSTDT**TETFRPGGGD
 MRDNWRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVSLGFLGAAGSTMGAAS
 MTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKLDTHWGIKQLQARVLAVEHYLRDQQLLGIW
 GCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQ
 DLLALD

(e) BG505_SOSIP_VLC1-03_DS21_mC

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLENVTEE
 FNMWKNMVEQMHTDIISLWDQSLKPCKLTPLCVTLQCTDWDNATLANMTGEIKNCSFNMT
 TELRDKKQKVYSLFYELDIPIENEYISNNNTSNTSYRLINCNTSAITQACPKVSFEPIPIH
 YCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEEVVMIRSENIT
 NNAKNILVQFNTVPVQINCTRPNNNTRKSIRIGPGCAFYATGDIIGDIRQAHCNVSATQWEQT
 LKGIAAKLLEHFGNNTIIRFANSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNGSSWNLNKTKE
 NTTNLENGTITLPCRICKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGGNKSAGIE
 TFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVFLGFLGAA
 GSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQQHLLKLTWVGIKQLQARVLAVERYLR
 DQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQ
 NQQEKNEQDLLALD

(f) BG505_SOSIP_VLC2-04_DS21_mC

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLENVTEE
 FNMWKNMVEQMHTDIISLWDQSLKPCKLTPLCVTLQCSDEYEGNTTRQNITMKEEKGEIKN
 CSFNMTTEL RDKKQKVYSLFYKLDITPIEEDNNSNNSSANSSNSNANYTNYRLINCNTSAI
 TQACPKVSFEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNG
 SLAEEVVMIRSENITNNAKNILVQFNTVPVQINCTRPNNNTRKSIRIGPGCAFYATGDIIGDI
 RQAHCNVSGTKWKNTLKQIVKKLGDHFGNNTIIRFANSSGGDLEVTTTHSFNCGGEFFYCNTS
 GLFNSTWNRTNGTWNDVEGLNYTNGNDTITLPCRICKQIINMWQRIGQAMYAPPIQGVIRCVS
 NITGLILTRDGGNDTDKNETFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRR
 RRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQQHLLKLT
 VWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQW
 DKEISNYTQIIYGLLEESQNQQEKNEQDLLALD

(g) BG505_SOSIP_VLC2-08_DS21_mC

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLENVTEE
 FNMWKNMVEQMHTDIISLWDQSLKPCKLTPLCVTLQCVTLKNCSNSNCSISRNISIEMDG
 EIKNCSFNMTTEL RDKKQKVYSLFYRLDIVPIESSNNSQLSNNSQVSNNQSNNYSQYRLIN
 CNTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVST
 QLLNGSLAEEVVMIRSENITNNAKNILVQFNTVPVQINCTRPNNNTRKSIRIGPGCAFYATG
 DIIGDIRQAHCNVSKKDWEKTLQQVATKLGQHFGNNTIIRFANSSGGDLEVTTTHSFNCGGEF
 FYCNTSGLFNSTWIRNSSNSTWNSSASNSTELNSNITLPCRICKQIINMWQRIGQAMYAPPIQ
 GVIRCVSNITGLILTRDGGHETENKTETFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRC

KRRVVGRRRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEA
 QQHLLKLTWVGIKQLQARVLAVERYLRDQQLLGIWGC SGKLI CCTNVPWNSSWSNRNLSEIW
 DNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALD

(h) BG505_SOSIP_VLC3-13_DS21_mC

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQEIHLENVTEE
 FNMWKNNMVEQMHTDIIISLWDQSLKPCCKLTPLCVTLQCTNAALTNVTITNGPNITEEIRNC
 SFNMTTELDRDKKQKVYSLFYKLDLVQINGSGGEYRLINCNTSAITQACPKVSFEPIPIHYCA
 PAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEEVVMIRSENITNNA
 KNILVQFNTVPVQINCTRPNNNTRKSIRIGPGCAFYATGDIIGDIRQAHCNVSGTKWNETLKQ
 VAGKL RKHFGNNTIIRFANSSGGDLEVTTHSFNCGGEFFYCNTSGLFNSTWPENGTMESNG
 TITLPCRICKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGGNSTTDTETFRPGGGD
 MRDNWRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVFLGFLGAAGSTMGAAS
 MTLTVQARNLLSGIVQQQSNLLRAPEAQQHLLKLTWVGIKQLQARVLAVERYLRDQQLLGIW
 GCSGKLI CCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQ
 DLLALD

(i) AC10_SOSIP_olio6_mC

AVEQTWVTVYYGVPVWKEANTTLFCASDAKAYNTEVHNWATHACVPTDPNPQEVELENVTE
 NFNMWKNNMVDQMHEDEIISLWDQSLKPCVKLTPLCVTLSCDNDVGNSTSTNNSRWDKMEKGE
 IKNC SFNITTNMRDKMQQYALFWKLDVVP IEEGKNNNSSFTDYRLISCNTSVITQACPKVT
 FEPIPIHYCAPAGFALLKCKDKKFNGTGPCKNVSTVQCTHGIKPVVSTQLLLNGSLAEEVVM
 IRSENFSSNARTIIVQLNTSVEIKCIA PNNFTVKG I HIGPGRAFYYMGDIIGDIRQAHCNIS
 RQNWNTLQKIAEKLREQFGNKTIVFRQSSGGDPEIVMHTFNCAGMFFYCNTAELFNSTWYA
 NGTISIGGKNKTNIILPCRICKLFINMWQEVGKAMYAPPISGQIRCSSNITGLLLTRDGGGRN
 QTDNQTEIFRPVGGDMKNNWRSELYKYKVVR I EPLGIAPTRCKRRVVGRRRRRRRAVGIGALS
 LGFLGAAGSTMGAASMTLTVQARLLL SGIVQQQNNLLRAPEPQQHLLQDTHWGIKQLQARVL
 AVEHYLRDQQLLGIWGC SGKLI CCTAVPWNVSWNNRSVDDI WENMTWMQWDREISNYTSLIY
 TLIEESQNQQEKNEQELLALD

(j) BG505_SOSIP_SET224_4_mC (BG505_SOSIP_R4_mC)

AENLWVTVYYGVPVWKDAETTLFCASDAKAYSTEKHNWATHACVPTDPNPQEVVLENVTEN
 FNMWKNNMVEQMHEDEIISLWDQSLKPCVKLTPLCVTLNCTELNDTNTTATNSSGRVIEDKEI
 KNCSFNMTTSLRDKVQRVYSLFNKFDIVPIDNSNDSYRLISCNTSAITQACPKVSFEPIPIH
 YCAPAGFAILKCN DKEFN GTGPKSVSTVQCTHGI R PVVSTQLLLNGSLAEEVVIIRSENFT
 NNAKTILVQLNEP VVINCTRPNNNTVKSIRIGPGQAFYYTGEIIGDIRQAHCTVSRETWNKT
 LGRVVEQLREQFRNKTII VFNQSSGGDPEIVMHSFNCGGEFFYCNSTQLFNSTWYGNETETG
 GTNDTIGNITLPCRICKQIINMWQEVGKAMYAPPIRGQISCSSNITGLILTRDGGNNNETNTT
 ETFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTCKRRVVGRRRRRRRAVGIGAMSLGFLGA
 AGSTMGAASLTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARVLAVEHYL

RDQQLGIWGCSGKLI CCTNVPWN TSWSNKSLDQIWDNMTWLEWDREISNYTQLIYNLLEES
QNQQEKNEQDLLALD

8. The non-naturally occurring protein of claim 1 or 2, wherein the trimer comprises any one of:

(a) BG505_SOSIP_MD39_congly_mC

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQEIHLENVTEE
FNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELKNCSFNMT
TEL RDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIH
YCAPAGFAILKCKDKKFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEVEIIRSENIT
NNAKNILVQLNTSVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNVSKATWNET
LGKVVVKQLRKHFNGNTIIRFAQSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQG
SNSTGSNDSITLPCRICKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTE
TFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRAVGIGAVSLGFLGAA
GSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARVLAVEHYLR
DQQLLGIWGCSGKLI CCTNVPWN SWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQ
NQQEKNEQDLLALD

(b) BG505_SOSIP_MD39_CtCys_mC

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQEIHLENVTEE
FNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELKNCSFNMT
TEL RDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIH
YCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEEVEIIRSENIT
NNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNVSKATWNET
LGKVVVKQLRKHFNGNTIIRFAQSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQG
SNSTGSNDSITLPCRICKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTE
TFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRAVGIGAVSLGFLGAA
GSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARVLAVEHYLR
DQQLLGIWGCSGKLI CCTNVPWN SWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQ
NQQEKNEQDLLALDGTKHHHHHHHC

(c) BG505_SOSIP_MD39_CD4KO4_mC

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQEIHLENVTEE
FNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELKNCSFNMT
TEL RDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIH
YCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEEVEIIRSENIT
NNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNVSKATWNET
LGKVVVKQLRKHFNGNTIIRFAQSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQG
SNSTGSNDSITLPCRICKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTE
TFRPGGTDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRAVGIGAVSLGFLGAA
GSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARVLAVEHYLR

DQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQ
NQQEKNEQDLLALD

(d) BG505_SOSIP_D664_olio6_CD4KO4_mC

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQEIHLENVTEE
FNMWKNNMVEQMHEIDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELKNCSFNMT
TELDRKKQKVYSLFWRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIH
YCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLNGSLAEEVVIIRSENIT
NNAKNILVQLNTPVQINCTAPNNFTVKSIRIGPGQAFYYMGDIIGDIRQAHCNVSKATWNET
LGKVVQQLRKHFNGNTIIRFAQSSGGDLEVTTTHSFNCGGMFFYCNTSGLFNSTWISNTSVQG
SNSTGSNDSITLPCRIKLIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTE
TFRPGGTDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVSLGFLGAA
GSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARVLAVEHYLR
DQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQ
NQQEKNEQDLLALD

(e) AC10_SOSIP_olio6_congly_mC

AVEQTWVTVYYGVPVWKEANTTLFCASDAKAYNTEVHNWATHACVPTDPNPQEVELENVTE
NFMWKNNMVDQMHEIDIISLWDQSLKPCVKLTPLCVTLSCDNDVGNSTSTNNSRWDKMEKGE
IKNCSFNITTNMRDKMQQYALFWKLDVVPIEEGKNNNSFTDYRLISCNTSVITQACPKVT
FEPIPIHYCAPAGFALLKCKDKKFNGTGPCKNVSTVQCTHGIKPVVSTQLLNGSLAEEVVI
IRSENFSSNNARTIIVQLNTSVEINCTAPNNFTVKGIIHIGPGRAFYYMGDIIGDIRQAHCNIS
RQNWNTTLKQIAEKLREQFGNKTIVFRQSSGGDPEIVMHTFNCAGMFFYCNTSELFNSTWYA
NGTISIGGGNKTNIILPCRIKLFINMWQEVGKAMYAPPISGQIRCSSNITGLLLLTRDGGGRGN
QTDNQTEIFRPVGGDMKNNNRSELYKYKVVRIEPLGIAPTRCKRRVVGRRRRRRRAVGIGALS
LGFLGAAGSTMGAASMTLTVQARLLLSGIVQQQNNLLRAPEPQQHLLQDTHWGIKQLQARVL
AVEHYLRDQQLLGIWGCSGKLICCTAVPWNVSWNNRSVDDIWENMTWMQWDREISNYTSLIY
TLIEESQNNQQEKNEQELLALD

9. The non-naturally occurring protein of claim 1 or 2, wherein the trimer comprises any one of:

(a) 191084_SOSIP_MD39

TENLWVTVYYGVPVWRDAETTLFCASDAKAYDTEMHNWATHACVPTDPNPQEIDLNVTEK
FNMWKNNMVEQMHEIDIISLWDQSLKPCVKLTPLCVTLNCTAITNDTRGNETGINRTVETTEM
TNCSFNMTTELDRDRKKVNALFYKLDIVQIGENSSSQYRLINCNTSVITQACPKVTFEPIPI
HYCAPAGFAILKCKDKKFNGTGTCTRVSSVQCTHGIKPVVSTQLLNGSLAEGQVVIIRSENI
SDNAKTIIVQLNESVPINCTRPNNNTVRGIHLGPGQTFYTDIIGDIRQAHCNVSESKWNKA
LQEVVKQLRQHWNKTIIFKSSSGGDLEITTHSFNCGGEFFYCNTSGLFNSTWNIAGNRNTNDT
KSNETITLPCRIKQIVNVWQRVGQAIYAPPIAGVIRCNSNITGLLLVRDGGATNNTDETFRP
GGGNMRDNWRSELYKYKVVKIEPLGVAPTRCRRRVVERRRRRRRAVGLGAVSIGFLGAAGSTM
GAASVTLTVQARQLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARVLAVEHYLKDQQL

LGIWGCSGKLICCTVPWNSSWSNKSQNEIWDNMTWLQWDKEISNYTQLIYSLIEESQNQQE
KNEQEELLALD

(b) 001428_SOSIP_MD39

VENLWVTVYYGVPVWKEARTTLFCASDAKAYETEVHNVWATHACVPTDPNPQEMVLGNVTEN
FNMWKNDMVDQMHEDEVISLWAQSLKPCVKLTPLCVTLECTQVNATQGNTTQVNVTQVNGDEM
KNCSFNNTTTEIRDKKQKAYALFYRLDLVPLERENRGDSNSASKYILINCNTSAITQACPKVN
FDPIPIHYCTPAGYAILKCNKTFNGTGSCNNVSTVQCTHGIKPVVSTQLLNGLSLAEFEEII
IRSENLTDNVKTIIVHLDQSVEIVCTRPNNNTVKSIRIGPGQTFYYTGDIIGNIREAHCNIS
EKKWHEMLRRVSEKLAEHFPNKTIKFTSSSGGDLEITTHSFNCRGEFFYCNTSGLFNSTYMP
NGTYMPNGTNNNSNTIILPCRIKQIINMWQEVGRAMYAPPIAGNITCNSNITGLLLVRDGGK
NNNTEIFRPGGGDMRDNRSELYKYKVVEIKPLGVAPTRCKRRVVGRRRRRRAVGLGAVSLG
FLGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLQAPEPQQHLLQDTHWGIKQLQTRVLAI
EHYLDQQLLGIWGCSGKLICCTAVPWNSSWSNKS LTDIWDNMTWMQWDREVSNYTGIIYRL
LEDSONQQERNEQDLLALD

(c) AC10_SOSIP_MD39

AVEQTVWTVYYGVPVWKEANTTLFCASDAKAYNTEVHNVWATHACVPTDPNPQEVELENVTE
NFMWKNNMVDQMHEDEIISLWDQSLKPCVKLTPLCVTLSCDNDVGNSTSTNNSRWDKMEKGE
IKNCSFNITTNMRDKMQQYALFYKLDVVPIEEGKNNNSFTDYRLISCNTSVITQACPKVT
FEPIPIHYCAPAGFALLKCKDKKFNGTGPCKNVSTVQCTHGIKPVVSTQLLNGLSLAEFEEV
IRSENFSSNNARTIIVQLNTSVEIKCIRPNNNTVKGIIHIGPGRFYTGDIIGDIRQAHCNIS
RQNWNNTLKQIAEKLREQFGNKTIIVFRQSSGGDPEIVMHTFNCAGEFFYCNTAELFNSTWYA
NGTISIGGGNKNTNIILPCRIKQFINMWQEVGKAMYAPPISGQIRCSSNITGLLLTRDGGRGN
QTDNQTEIFRPVGGDMKNNNRSELYKYKVVRIEPLGIAPTRCKRRVVGRRRRRRAVGIGALS
LGFLGAAGSTMGAASMTLTVQARLLLSGIVQQQNNLLRAPEPQQHLLQDTHWGIKQLQARVL
AVEHYLRDQQLLGIWGCSGKLICCTAVPWNVSWNNRSVDDIWENMTWMQWDREISNYTSLIY
TLIEESQNQQEKNEQEELLALD

(d) AC10_SOSIP_olio6

AVEQTVWTVYYGVPVWKEANTTLFCASDAKAYNTEVHNVWATHACVPTDPNPQEVELENVTE
NFMWKNNMVDQMHEDEIISLWDQSLKPCVKLTPLCVTLSCDNDVGNSTSTNNSRWDKMEKGE
IKNCSFNITTNMRDKMQQYALFWKLDVVPIEEGKNNNSFTDYRLISCNTSVITQACPKVT
FEPIPIHYCAPAGFALLKCKDKKFNGTGPCKNVSTVQCTHGIKPVVSTQLLNGLSLAEFEEV
IRSENFSSNNARTIIVQLNTSVEIKCIAPNNFTVKGIIHIGPGRFYMGDIIGDIRQAHCNIS
RQNWNNTLKQIAEKLREQFGNKTIIVFRQSSGGDPEIVMHTFNCAGMFFYCNTAELFNSTWYA
NGTISIGGGNKNTNIILPCRIKLFINMWQEVGKAMYAPPISGQIRCSSNITGLLLTRDGGRGN
QTDNQTEIFRPVGGDMKNNNRSELYKYKVVRIEPLGIAPTRCKRRVVGRRRRRRAVGIGALS
LGFLGAAGSTMGAASMTLTVQARLLLSGIVQQQNNLLRAPEPQQHLLQDTHWGIKQLQARVL
AVEHYLRDQQLLGIWGCSGKLICCTAVPWNVSWNNRSVDDIWENMTWMQWDREISNYTSLIY
TLIEESQNQQEKNEQEELLALD

(e) ZM197M_MD39

MEQLWVTVYYYGVPVWKEAKATLFCASDAKAYEKEVHNWATHACVPTDPNPQEIPLGNVTEN
 FNMWKNDMADQMHEDIISLWDQSLKPCVKLTPLCVTLNCSDATSNITTKNATNTNTTSTDNRN
 ATSNDETEMKGEIKNCTFNITTEVRDRKTKQRALFYKLDVVPLEEEKNSSSKNSSYKEYRLIS
 CNTSTITQACPKVSFDPIPIHYCAPAGYAILKCNKTFNGTGPCHNVTSTVQCTHGIKPVVST
 QLLNGSLAEEEEIIIRSENLTDNKTIIIVHLNESVEINCTRPNNNTVKSVRIGPGQTFYTG
 EIIIGDIRQAHCNLSKSNWTTTLKRIEKKLKEHFNNATIKFESSAGGDLEITTHSFNCRGEFF
 YCNTSGLFNSSLNDTDGTSNSTSNATITLPCRIKQIINMWQEVGRAMYASPIAGIITCKSN
 ITGLLLTRDGGNKSAGIETFRPGGGNMKDNWRSELYKYKVVEIKPLGIAPTSCKRRVVERRR
 RRRRAAGIGAVSLGFLGAAGSTMGAASVMLTVQARQLLSGIVQQQSNLLRAPEPQQHMLQDTH
 WGIKQLQTRVLAIEHYLKDQQLLGLWGCSGKLICTAVPWNTSWSNKSDEIWDNMTWMQWD
 REIDNYTQVIYQLLEVSQNNQEQKNENDLLALD

(f) B41_SOSIP_D664_MD39

AAKKWVTVYYYGVPVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNPQEIVLGNVTEN
 FNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCNNVNTNNTNNSNATISDWEKME
 TGEMKNCSFNVTTSIRDKIKKEYALFYKLDVVPLENKNNINNTNITNYRLINCNTSVITQAC
 PKVSFEPIPIHYCAPAGFAILKCNKTFNGSGPCTNVSTVQCTHGIRPVVSTQLLNGSLAE
 EEIVIRSENITDNAKTIIIVQLNEAVEINCTRPNNNTVKSIIHIGPGRAFYYTGDIIGNIRQAH
 CNISKARWNETLGQIVAKLEEQFPNKTIIFNHSSGGDPEIVTHSFNCGGEFFYCNTTPLFNS
 TWNNTRTDDYPTGGEQNITLQCRIKQIINMWQGVGKAMYAPPIRGQIRCSSNITGLLLTRDG
 GRDQNGTETFRPGGGNMRDNWRSELYKYKVVKIEPLGIAPTACKRRVVQRRRRRRRAVGLGAF
 SLGFLGAAGSTMGAASMLTVQARLLLSGIVQQQNNLLRAPEPQQHMLQDTHWGIKQLQARV
 LAVEHYLRDQQLLGIWGCSGKLICTNVPWNDSWSNKTINEIWDNMTWMQWEKEIDNYTQHI
 YTLLEVSQIQEQKNEQELLELD

10. The non-naturally occurring protein of claim 1 or 2, wherein the trimer comprises any one of:

(a) BG505_SOSIP_MD39_link14

AENLWVTVYYYGVPVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQEIHLENVTEE
 FNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNVNTNITDDMRGELKNCSFNMT
 TELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIH
 YCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLNGSLAEEVEIIIRSENIT
 NNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNVSKATWNET
 LGKVVKQLRKHFGNNTIIRFAQSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQG
 SNSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTE
 TFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGSHSGSGSGSGGHA AVGIGAV
 SLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARV
 LAVEHYLRDQQLLGIWGCSGKLICTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQII
 YGLLEESQNNQEQKNEQDLLALD

(b) BG505_SOSIP_MD39_CP1.1

VSLGFLGAAGSTMGAAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQAR
 VLAVEHYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQI
 IYGLLEESQNQQEKNEQDLLALDGGSGSGGGSGSGSSAENLWVTVYYGVPVWKDAETTLFC
 ASDAKAYETEKHNVWATHACVPTDPNPQEIHLNVTEEFNMWKNNMVEQMHEDIISLWDQSL
 KPCVKLTPLCVTLQCTNVTNNITDDMRGELKNCSFNMTTEL RDKKQKVYSLFYRLDVVQINE
 NQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDKKFN GTGPCP
 SVSTVQCTHGIKPVVSTQLLLNGSLAEEEEVIIRSENITNNAKNILVQLNTPVQINCTRPNNN
 TVKSIRIGPGQAFYYTGDIIGDIRQAHCNVSKATWNETLGKVVKQLRKHFNGNTIIRFAQSS
 GGDLEVTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQGSNSTGSNDSITLPCRIKQIINMWQ
 RIGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGGGDMRDNRSELYKYKVVK
 IEPLGVAPTRCKRRVVG

(c) BG505_SOSIP_MD39_CP1.2

VSLGFLGAAGSTMGAAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQAR
 VLAVEHYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQI
 IYGLLEESQNQQEKNEQDLGGSGSGGGSGSGSSGSGLWVTVYYGVPVWKDAETTLFCASDA
 KAYETEKHNVWATHACVPTDPNPQEIHLNVTEEFNMWKNNMVEQMHEDIISLWDQSLKPCV
 KLTPLCVTLQCTNVTNNITDDMRGELKNCSFNMTTEL RDKKQKVYSLFYRLDVVQINENQGN
 RSNNSNKEYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDKKFN GTGPCPSVST
 VQCTHGIKPVVSTQLLLNGSLAEEEEVIIRSENITNNAKNILVQLNTPVQINCTRPNNNTVKS
 IRIGPGQAFYYTGDIIGDIRQAHCNVSKATWNETLGKVVKQLRKHFNGNTIIRFAQSSGGDL
 EVTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQGSNSTGSNDSITLPCRIKQIINMWQRIGQ
 AMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGGGDMRDNRSELYKYKVVKIEPL
 GVAPTRCKRR

(d) BG505_SOSIP_MD39_CP2

VSLGFLGAAGSTMGAAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQAR
 VLAVEHYLRDQQLLGIWGCSGKLICCTGGSGVTVYYGVPVWKDAETTLFCASDAKAYETEKH
NVWATHACVPTDPNPQEIHLNVTEEFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVT
 LQCTNVTNNITDDMRGELKNCSFNMTTEL RDKKQKVYSLFYRLDVVQINENQGNRSNNSNKE
 YRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDKKFN GTGPCPSVSTVQCTHGIK
 PVVSTQLLLNGSLAEEEEVIIRSENITNNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPGQA
 FYYTGDIIGDIRQAHCNVSKATWNETLGKVVKQLRKHFNGNTIIRFAQSSGGDLEVTTHSFN
 CGGEFFYCNTSGLFNSTWISNTSVQGSNSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQ
 GVIRCVSNITGLILTRDGGSTNSTTETFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCK
 RGGSGSGVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLA
 LD

(e) BG505_SOSIP_MD39_CP3

VSLGFLGAAGSTTMGAA~~SMT~~TLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQAR
 VLAVEHYLRDQQLLGIWGC~~SG~~KLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQI
 IYGLLEESGGSGSGAGGLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNWATHACVPTD
 PNPQEIHLENVTEEFNMWKNNMVEQM~~HEDI~~ISLWDQSLKPCVKLTPLCVTLQCTNVTNNITD
 DMRGELKNCSFNMTTEL~~RDKKQKVYSLFYRLDVVQINENQGNRSNNSN~~KEYRLINCNTSAIT
 QACPKVSFEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGS
 LAEEEEVIIRSENITNNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIR
 QAHCNVSKATWNETLGKVVKQLRKHF~~GNNT~~IIRFAQSSGGDLEVTT~~HSFNC~~GGEFFYCNTSG
 LFNSTWISNTSVQGSNSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQGVIRCVSNITGL
 ILTRDGGSTNSTTETFRPGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGGSGSGQ
 NQQEKNEQDLLALD

11. The non-naturally occurring protein of claim 1 or 2, wherein the trimer comprises any one of:

(a) BG505_MD39_GRSF4

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNSSEIHLENVTE
 EFNMWKNMVEQM~~HEDI~~ISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELKNCSFN
 MTTEL~~RDKKQKVYSLFYRLDVVQINENQGNRSNNSN~~KEYRLINCNTSAITQACPKVSFEPI
 PIHYCAPAGFAILKCKDKKFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEEVIIRS
 ENITNNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNVSKA
 TWNETLGKVVKQLRKHF~~GNNT~~IIRFAQSSGGDLEVTT~~HSFNC~~GGEFFYCNTSGLFNSTWIS
 NTSVQGSNSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGG
 STNSTTETFRPGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAV
 SLGFLGAAGSTMGAA~~SMT~~TLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQAR
 VLAVEHYLRDQQLLGIWGC~~SG~~KLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQ
 IYGLLEESQNQEKNNQSLALDNGS

(b) BG505_MD39_GRSF7

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNSSEIHLENVTE
 EFNMWKNMVEQM~~HEDI~~ISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELKNCSFN
 MTTEL~~RDKKQKVYSLFYRLDVVQINENQGNRSNNSN~~KEYRLINCNTSAITQACPKVSFEPI
 PIHYCAPAGFAILKCKDKKFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEEVIIRS
 ENITNNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNVSKA
 TWNETLGKVVKQLRKHF~~GNNT~~IIRFAQSSGGDLEVTT~~HSFNC~~GGEFFYCNTSGLFNSTWIS
 NTSVQGSNSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGG
 STNSTTETFRPGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAV
 SLGFLGAAGSTMGAA~~SMT~~TLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQAR
 VLAVEHYLRDQQLLGIWGC~~SG~~KLICCTNVPWNSSWSNANLSEIWDNMTWLQWDKNISNYTQ
 IYGLLEESQNQEKNNQSLALDNGS

(c) BG505_MD39_CP1.2_GRSF4

VSLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQA
 RVLAVEHYLRDQQLLGIWGCSGKLI CCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYT
 QIIYGLLEESQNQQEKNEQDL**GNGSGGGSGSGGSSGL**WVTVYYGVPVWKDAETTLFCA
 SDAKAYETEKHNWATHACVPTDPN**SSEI**HLENVTEEFNMWKNNMVEQMHEDIISLWDQSL
 KPCVKLTPLCVTLQCTNVTNNITDDMRGELKNCSFNMTTEL RDKKQKVYSLFYRLDVVQIN
 ENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDKKFNGTG
 P**QNV**STVQCTHGIKPVVSTQLLLNGSLAEEVVIIRSENITNNAKNILVQLNT**SVQ**INCTRP
 NNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNVSKATWNETLGKVVQQLRKHFNGNTIIRF
 AQSSGGDLEVTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQGSNSTGSNDSITLPCRICKQI
 INMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGGDMRDNRSELY
 KYKVVKIEPLGVAPTRCKRR

(d) BG505_MD39_link14_GRSF4

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPN**SSEI**HLENVTE
 EFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELKNCSFN
 MTTEL RDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPI
 PIHYCAPAGFAILKCKDKKFNGTG**PCQNV**STVQCTHGIKPVVSTQLLLNGSLAEEVVIIRS
 ENITNNAKNILVQLNT**SVQ**INCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNVSKA
 TWNETLGKVVQQLRKHFNGNTIIRFAQSSGGDLEVTTHSFNCGGEFFYCNTSGLFNSTWIS
 NTSVQGSNSTGSNDSITLPCRICKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGG
 STNSTTETFRPGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVG**SHSGSGSGSGGGH**
AAVGIGAVSLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHW
GIKQLQARVLAVEHYLRDQQLLGIWGCSGKLI CCTNVPWNSSWSNRNLSEIWDNMTWLQWD
KEISNYTQIIYGLLEESQNQQEKNNQSLALDNGS

(e) BG505_MD39_link14_GRSF7

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPN**SSEI**HLENVTE
 EFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELKNCSFN
 MTTEL RDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPI
 PIHYCAPAGFAILKCKDKKFNGTG**PCQNV**STVQCTHGIKPVVSTQLLLNGSLAEEVVIIRS
 ENITNNAKNILVQLNT**SVQ**INCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNVSKA
 TWNETLGKVVQQLRKHFNGNTIIRFAQSSGGDLEVTTHSFNCGGEFFYCNTSGLFNSTWIS
 NTSVQGSNSTGSNDSITLPCRICKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGG
 STNSTTETFRPGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVG**SHSGSGSGSGGGH**
AAVGIGAVSLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHW
GIKQLQARVLAVEHYLRDQQLLGIWGCSGKLI CCTNVPWNSSWSNANLSEIWDNMTWLQWD
KNISNYTQIIYGLLEESQNQQEKNNQSLALDNGS

12. The non-naturally occurring protein of claim 1 or 2, wherein the trimer comprises any one of:

(a) BG505_MD39_CP1.2_GRSF4_qLoops1

VSLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQAR
VLAVEHYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQI
IYGLLEESQNQQEKNEQDL**GGNGSGGGSGSGGNGSS**GLWVTVYYGVPVWKDAETTLFCASDA
KAYETEKHNVWATHACVPTDPN**SSEI**HLENVT EEFNMWKNMVEQM HEDIISLWDQSLKPCV
KLTPLCVTLQCTNVTNNITD**NMT**GELKNCSFNMTTEL RDKKQKVYSLFYRLDVVQIN**NGSGGE**
YRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDKKFNGTGPC**QNV**STVQCTHGIK
PVVSTQLLLNGSLAEEVVIIRSENITNNAKNILVQLNT**SVQ**INCTRPNNNTVKSIRIGPGQA
FYTGDIIGDIRQAHCNVSKATWNETLGKVVQQLRKHFNGNTIIRFAQSSGGDLEVTTHSFN
CGGEFFYCNTSGLFNSTW**PENGTMEGSNGT**ITLPCRIKQIINMWQRIGQAMYAPPIQGVIRC
VSNITGLILTRDGGSTNSTTETFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRR

(b) BG505_MD39_CP1.2_GRSF7_qLoops1

GGNSSSGSLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIK
QLQARVLAVEHYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEIS
NYTQIIYGLLEESQN**QNES**NEQDL**GGNGSGGGSGSGGNGSS**GLWVTVYYGVPVWKDAETTLF
CASDAKAYETEKHNVWATHACVPTDPN**SSEI**HLENVT EEFNMWKNMVEQM HEDIISLWDQS
LKPCVKLTPLCVTLQCTNVTNNITD**NMT**GELKNCSFNMTTEL RDKKQKVYSLFYRLDVVQIN
GSGGEYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDKKFNGTGPC**QNV**STVQC
THGIKPVVSTQLLLNGSLAEEVVIIRSENITNNAKNILVQLNT**SVQ**INCTRPNNNTVKSIRI
GPGQAFYYTGDIIGDIRQAHCNVSKATWNETLGKVVQQLRKHFNGNTIIRFAQSSGGDLEVT
THSFNCGGEFFYCNTSGLFNSTW**PENGTMEGSNGT**ITLPCRIKQIINMWQRIGQAMYAPPIQ
GVIRCVSNITGLILTRDGGSTNSTTETFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRC**N**
RS

(c) BG505_MD39_GRSF4_qLoops1

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPN**SSEI**HLENVT EEFNMWKNMVEQM HEDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITD**NMT**GELKNCSFNMT
TEL RDKKQKVYSLFYRLDVVQIN**NGSG**GGEYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFA
ILKCKDKKFNGTGPC**QNV**STVQCTHGIKPVVSTQLLLNGSLAEEVVIIRSENITNNAKNILV
QLNTSVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNVSKATWNETLGKVVQQL
RKHFNGNTIIRFAQSSGGDLEVTTHSFNCGGEFFYCNTSGLFNSTW**PENGTMEGSNGT**ITLPC
CRIKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGGGDMRDNR
SELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVSLGFLGAAGSTMGAASMTLTV
QARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARVLAVEHYLRDQQLLGIWGCSGK
LICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEK**NNQ**SLAL
DNGS

13. The non-naturally occurring protein of claim 1 or 2, wherein the trimer comprises any one of:

(a) BG505_SOSIP_MD39_2JD6

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLENVTEE
 FNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELKNCSFNMT
 TELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIH
 YCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEEEEVIIRSENIT
 NNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNVSKATWNET
 LGKVVKQLRKHFNGNTIIRFAQSSGGDLEVTTSHFNCGGEFFYCNTSGLFNSTWISNTSVQG
 SNSTGSNDSITLPCRICKIINMWQRIGQAMYAPPIQGIVRCVSNITGLILTRDGGSTNSTTE
 TFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVSLGFLGAA
 GSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARVLAVEHYLR
 DQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQ
 NQQEKNEQDLLALD**GS**GGLSERMLKALNDQLNRELYSAYLYFAMAAYFEDLGLEGFANWMKA
 QAEEEIGHALRFYNYIYDKNGRVELDEIPKPPKEWESPLKAFAEAYEHEKFISKSIYELAAL
 AEEKDYSTRAFLEWFINEQVEEEASVKKILDKLKFAKDSPQILFMLDKELSARAPKLPGLL
 MQGGE**

(b) BG505 SOSIP MD39 E2p (also referred to as "MD39-1b5s")

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLENVTEE
 FNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELKNCSFNMT
 TELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIH
 YCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEEEEVIIRSENIT
 NNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNVSKATWNET
 LGKVVKQLRKHFNGNTIIRFAQSSGGDLEVTTSHFNCGGEFFYCNTSGLFNSTWISNTSVQG
 SNSTGSNDSITLPCRICKIINMWQRIGQAMYAPPIQGIVRCVSNITGLILTRDGGSTNSTTE
 TFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVSLGFLGAA
 GSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARVLAVEHYLR
 DQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQ
 NQQEKNEQDLLALD**AS**GAAAKPATTEGEFPETREKMSGIRRAIAKAMVHSHKHTAPHVTLMDE
 ADVTKLVAHRKKFKAIAAEKGIKLTFLPVVKALVSALREYPVLNT [C/A/T] IDDETEEII
 QKHYYNIGIAADTDGRLLVPVIKHADRKPIFALAQEINELA EKARDGKLT PGEMKGASCTIT
 NIGSAGGQWFPTVINHPEVAILGIGRIA EKPIVRDGEIVAAPMLALSLSFDHRMIDGATAQK
 ALNHIKRLLSDPELLLM**

14. The non-naturally occurring protein of claim 1 or 2, wherein the trimer comprises any one of:

(a) BG505_MD39_gp160_dCT

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLENVTEE
 FNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELKNCSFNMT
 TELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIH
 YCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEEEEVIIRSENIT
 NNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNVSKATWNET

LGKVVKQLRKHFNGNTIIRFAQSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQG
 SNSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQGVIKCVSNITGLILTRDGGSTNSTTE
 TFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVSLGFLGAA
 GSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARVLAVEHYLR
 DQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQ
 NQQEKNEQDLLALDKWASLWNWFDISNWLWYIKIFIMIVGGLIGLRIVFAVLSVIHRVRQGY
 SPLS**

(b) BG505_MD39_gp160_dCT_GRSF4.1

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNSSEIHLENVTEE
 FNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELKNCSFNMT
 TELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPPIPIH
 YCAPAGFAILKCKDKKFNGTGPCQNVSTVQCTHGIKPVVSTQLLNGSLAEEVEIIRSENIT
 NNAKNILVQLNTSVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNVSKATWNET
 LGKVVKQLRKHFNGNTIIRFAQSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQG
 SNSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQGVIKCVSNITGLILTRDGGSTNSTTE
 TFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVSLGFLGAA
 GSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARVLAVEHYLR
 DQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQ
 NQQEKNEQDLLALDKWASLWNWFDISNWLWYIKIFIMIVGGLIGLRIVFAVLSVIHRVRQGY
 SPLS**

(c) BG505_MD39_gp160_GRSF4.1_m

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNSSEIHLENVTEE
 FNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELKNCSFNMT
 TELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPPIPIH
 YCAPAGFAILKCKDKKFNGTGPCQNVSTVQCTHGIKPVVSTQLLNGSLAEEVEIIRSENIT
 NNAKNILVQLNTSVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNVSKATWNET
 LGKVVKQLRKHFNGNTIIRFAQSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQG
 SNSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQGVIKCVSNITGLILTRDGGSTNSTTE
 TFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVSLGFLGAA
 GSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARVLAVEHYLR
 DQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQ
 NQQEKNEQDLLALDKWASLWNWFDISNWLWYIKIFIMIVGGLIGLRIVFAVLSVIHRVRQGY
 SPLSFQHTPNPRGLDRPERIEEDGEQDRGRSTRLVSGFLALAWDDLRLSLCLFCYHRLRDF
 ILIAARIVELLGHSSSLKGLRLGWEGLKYLWNLLAYWGRELKISAINLFDITIAIAVAEWTRV
 IEIGQRLCRAFLHIPRRIRQGLERALL**

(d) BG505_MD39_gp140-PDGFR

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLENVTEE
 FNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELKNCSFNMT

TEL RDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIH
 YCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAE EEEVIIRSENIT
 NNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAH CNVSKATWNET
 LGKVVKQLRKHF GNNTIIRFAQSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQG
 SNSTGSNDSITLPCR IKQIINMWQRIGQAMYAPPIQG VIRCVS NITGLILTRDGGSTNSTTE
 TFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRAVGIGAVSLGFLGAA
 GSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARVLAVEHYLR
 DQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQ
 NQQEKNEQDLLALDGGGSGGSGGSEQKLISEEDLGGSGGSGGSNAVGQDTQEVI VVPHSLPF
 KVVVISAILALVVLTIISLIILIMLWQKKPR**

(e) BG505_MD39_gp140-PDGFR_GRSF4.1_m

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPN **S**SEIHLENVTEE
 FNMWKNMVEQM HEDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELKNC SFNMT
 TEL RDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIH
 YCAPAGFAILKCKDKKFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAE EEEVIIRSENIT
 NNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAH CNVSKATWNET
 LGKVVKQLRKHF GNNTIIRFAQSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQG
 SNSTGSNDSITLPCR IKQIINMWQRIGQAMYAPPIQG VIRCVS NITGLILTRDGGSTNSTTE
 TFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRAVGIGAVSLGFLGAA
 GSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARVLAVEHYLR
 DQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQ
 NQQEKNEQDLLALDGGGSGGSGGSEQKLISEEDLGGSGGSGGSNAVGQDTQEVI VVPHSLPF
 KVVVISAILALVVLTIISLIILIMLWQKKPR**

(f) BG505_MD39_gp140-PDGFR_link14

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPN **P**QEIHLENVTEE
 FNMWKNMVEQM HEDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELKNC SFNMT
 TEL RDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIH
 YCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAE EEEVIIRSENIT
 NNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAH CNVSKATWNET
 LGKVVKQLRKHF GNNTIIRFAQSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQG
 SNSTGSNDSITLPCR IKQIINMWQRIGQAMYAPPIQG VIRCVS NITGLILTRDGGSTNSTTE
 TFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGSHSGSGGSGGHA AVGIGAV
 SLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARV
 LAVEHYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQII
 YGLLEESQNQQEKNEQDLLALDGGGSGGSGGSEQKLISEEDLGGSGGSGGSNAVGQDTQEVI
 VVPHSLPFKVVVISAILALVVLTIISLIILIMLWQKKPR**

(g) BG505_MD39_gp160-dCT_link14

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLENVTEE
 FNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELKNCSFNMT
 TELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIH
 YCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLNGSLAEEEEVIIRSENIT
 NNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNVSKATWNET
 LGKVVKQLRKHFNGNTIIRFAQSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQG
 SNSTGSNDSITLPCRICKQIINMWQRIGQAMYAPPIQGIVRCVSNITGLILTRDGGSTNSTTE
 TFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGSHSGSGSGSGGHAAVGIGAV
 SLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARV
 LAVEHYLRDQQLLGIWGCSGKLI CCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQII
 YGLLEESQNQQEKNEQDLLALDKWASLWNWFDISNWLWYIKIFIMIVGGLIGLRIVFAVLSV
 IHRVRQGYSPLS**

(h) BG505_MD39_gp160-dCT_link14_GRSF4.1

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPN**SS**EIHLENVTEE
 FNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELKNCSFNMT
 TELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIH
 YCAPAGFAILKCKDKKFNGTGPC**Q**NVSTVQCTHGIKPVVSTQLLNGSLAEEEEVIIRSENIT
 NNAKNILVQLNT**SV**QINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNVSKATWNET
 LGKVVKQLRKHFNGNTIIRFAQSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQG
 SNSTGSNDSITLPCRICKQIINMWQRIGQAMYAPPIQGIVRCVSNITGLILTRDGGSTNSTTE
 TFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGSHSGSGSGSGGHAAVGIGAV
 SLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARV
 LAVEHYLRDQQLLGIWGCSGKLI CCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQII
 YGLLEESQNQQEKNEQDLLALDKWASLWNWFDISNWLWYIKIFIMIVGGLIGLRIVFAVLSV
 IHRVRQGYSPLS**

(i) BG505_MD39_gp140-PDGFR_link14_GRSF4.1

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPN**SS**EIHLENVTEE
 FNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELKNCSFNMT
 TELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIH
 YCAPAGFAILKCKDKKFNGTGPC**Q**NVSTVQCTHGIKPVVSTQLLNGSLAEEEEVIIRSENIT
 NNAKNILVQLNT**SV**QINCTRPNNNTVKSIRIGPGQAFYYTGDIIGDIRQAHCNVSKATWNET
 LGKVVKQLRKHFNGNTIIRFAQSSGGDLEVTTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQG
 SNSTGSNDSITLPCRICKQIINMWQRIGQAMYAPPIQGIVRCVSNITGLILTRDGGSTNSTTE
 TFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCKRRVVGSHSGSGSGSGGHAAVGIGAV
 SLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQARV
 LAVEHYLRDQQLLGIWGCSGKLI CCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQII
 YGLLEESQNQQEKNEQDLLALDGGSGSGSGGSEQKLI SEEDLGGSGSGSGSNAVGQDTQEVI
 VVPHSLPFBKVVVISAILALVVLTIISLIILIMLWQKKPR**

(j) BG505_MD39_gp140-PDGFR_CP1.2_GRSF4.0

VSLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQAR
 VLAVEHYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQI
 IYGLLEESQNQQEKNEQDLGGNGSGGGSGSGGNGSSGLWVTVYYGVPVWKDAETTLFCASDA
 KAYETEKHNWATHACVPTDPNSSEIHLENVTEEFNMWKNMVEQMHEDIISLWDQSLKPCV
 KLTPLCVTLQCTNVTNNITDDMRGELKNCSFNMTTEL RDKKQKVYSLFYRLDVVQINENQGN
 RSNNNSNKEYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDKKFNGTGPCQNVST
 VQCTHGIKPVVSTQQLLNGSLAEEEVIIIRSENITNNAKNILVQLNTSVQINCTRPNNNTVKS
 IRIGPGQAFYYTGDIIGDIRQAHCNVSKATWNETLGKVVKQLRKHFNGNTIIRFAQSSGGDL
 EVTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQGSNSTGSNDSITLPCRIKQIINMWQRIGQ
 AMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGGGDMRDNRSELYKYKVVKIEPL
 GVAPTRCKRRGGSGSGSGGSEQKLISEEDLGGSGSGSGSNAVGGDTQEVIVPHSLPFKVVV
 ISAILALVVLTIISLIILIMLWQKKPR**

(k) BG505_MD39_gp160-dCT_CP2_GRSF4.1

VSLGFLGAAGSTMGAAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQAR
 VLAVEHYLRDQQLLGIWGCSGKLICCTGGSGVTVYYGVPVWKDAETTLFCASDAKAYETEKH
 NVWATHACVPTDPNSSEIHLENVTEEFNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVT
 LQCTNVTNNITDDMRGELKNCSFNMTTEL RDKKQKVYSLFYRLDVVQINENQGNRSNNNSKE
 YRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDKKFNGTGPCQNVSTVQCTHGIK
 PVVSTQQLLNGSLAEEEVIIIRSENITNNAKNILVQLNTSVQINCTRPNNNTVKSIRIGPGQA
 FYYTGDIIGDIRQAHCNVSKATWNETLGKVVKQLRKHFNGNTIIRFAQSSGGDLEVTTHSFN
 CGGEFFYCNTSGLFNSTWISNTSVQGSNSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQ
 GVIRCVSNITGLILTRDGGSTNSTTETFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCK
 RGGSGSGVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQI IYGLLEESQNQQEKNEQDLA
 LDKWASLWNWFDISNWLWYIKIFIMIVGGLIGLRIVFAVLSVIHRVRQGYSPLS**

(l) BG505_MD39_gp140-PDGFR_CP2_GRSF4.2

VSLGFLGAAGSTMGAAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQAR
 VLAVEHYLRDQQLLGIWGCSGKLICCTGGSGVTVYYGVPVWKDAETTLFCASDAKAYETEKH
 NVWATHACVPTDPNSSEIHLENVTEEFNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVT
 LQCTNVTNNITDDMRGELKNCSFNMTTEL RDKKQKVYSLFYRLDVVQINENQGNRSNNNSKE
 YRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDKKFNGTGPCQNVSTVQCTHGIK
 PVVSTQQLLNGSLAEEEVIIIRSENITNNAKNILVQLNTSVQINCTRPNNNTVKSIRIGPGQA
 FYYTGDIIGDIRQAHCNVSKATWNETLGKVVKQLRKHFNGNTIIRFAQSSGGDLEVTTHSFN
 CGGEFFYCNTSGLFNSTWISNTSVQGSNSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQ
 GVIRCVSNITGLILTRDGGSTNSTTETFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTRCK
 RGGSGSGVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQI IYGLLEESQNQQEKNNQSLLA
 LDNGSGGGSGGGSGGSEQKLISEEDLGGSGSGSGSNAVGGDTQEVIVPHSLPFKVVVISAIL
 ALVVLTIISLIILIMLWQKKPR**

(m) BG505_MD39_gp160_CP2_GRSF4.1

VSLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGIKQLQAR
VLAVEHYLRDQQLLGIWGCSGKLICCTGSGSVTVYYGVPVWKDAETTLFCASDAKAYETEKH
NVWATHACVPTDPNSSEIHLENVTEEFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVT
LQCTNVTNNITDDMRGELKNCSFNMTTELDRKKQKVYSLFYRLDVVQINENQGNRSNNSNKE
YRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDKKFNGTGPCQNVSTVQCTHGIK
PVVSTQLLLNGSLAEEEEVIIRSENITNNAKNILVQLNTSVQINCTRPNNNTVKSIRIGPGQA
FYYTGDIIGDIRQAHCNVSKATWNETLGKVVKQLRKHFGNNTIIRFAQSSGGDLEVTTHSFN
CGGEFFYCNTSGLFNSTWISNTSVQGSNSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQ
GVIRCVSNITGLILTRDGGSTNSTTETFRPGGDMRDNRSELYKYKVVKIEPLGVAPTRCK
RGGSGSGVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLA
LDKASLWNWFDISNWLWYIKIFIMIVGGLIGLRIVFAVLSVIHRVRQGYSPLSFQTHTPNP
RGLDRPERIEEEDGEQDRGRSTRLVSGFLALAWDDLRLCLFCYHRLRDFILIAARIVELLG
HSSLKGLRLGWEGLKYLWNLLAYWGRELKISAINLFDITIAIAVAEWTDREVIEIGQRLCRAFL
HIPRRIRQGLERALL**

15. The non-naturally occurring protein of claim 1 comprising a gp120 with a PGT121-class germline-targeting mutation, wherein the protein comprises any one of a protein comprising the sequence of any one of FIGS. 9-11 or 14.

16. A protein having at least 90% homology or identity with the sequence of the protein of any one of claims 2 to 15.

17. A protein having at least 95% homology or identity with the sequence of the protein of any one of claims 2 to 16.

18. A monomeric protein of any one of claims 2 to 15.

19. The protein of any one of claims 2-18 further comprising a tag for purification or biotinylation.

20. The protein of claim 19 wherein the tag for purification is a his tag.

21. The protein of claim 19 wherein the tag for biotinylation is an avi-tag.

22. The protein of any one of claims 2-21 further comprising an additional cysteine.

23. The protein of any one of claims any one of claims 2-22 fused to a multimerization motif.

24. A nucleic acid encoding the protein of any one of claims 2 to 23.

25. A nucleic acid having at least 90% homology or identity with the sequence of the nucleic acid of claim 24.

26. A nucleic acid having at least 95% homology or identity with the sequence of the nucleic acid of claim 24.

27. The nucleic acid of any one of claims 23-26 wherein the nucleic acid is a mRNA.

28. A method for eliciting an immune response comprising systemically administering to an animal in need thereof an effective amount of the protein of any one of claims 2-23.

29. The method of claim 28, wherein the animal is a mammal.

30. The method of claim 29, wherein the mammal is a human.

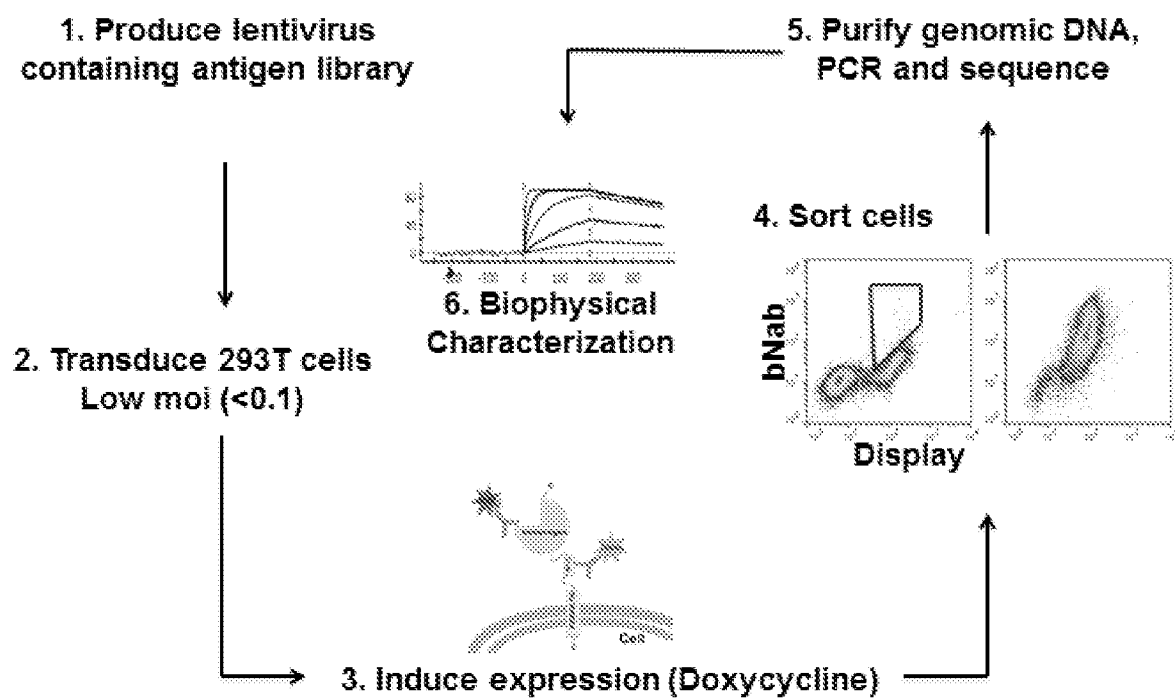


FIG. 1

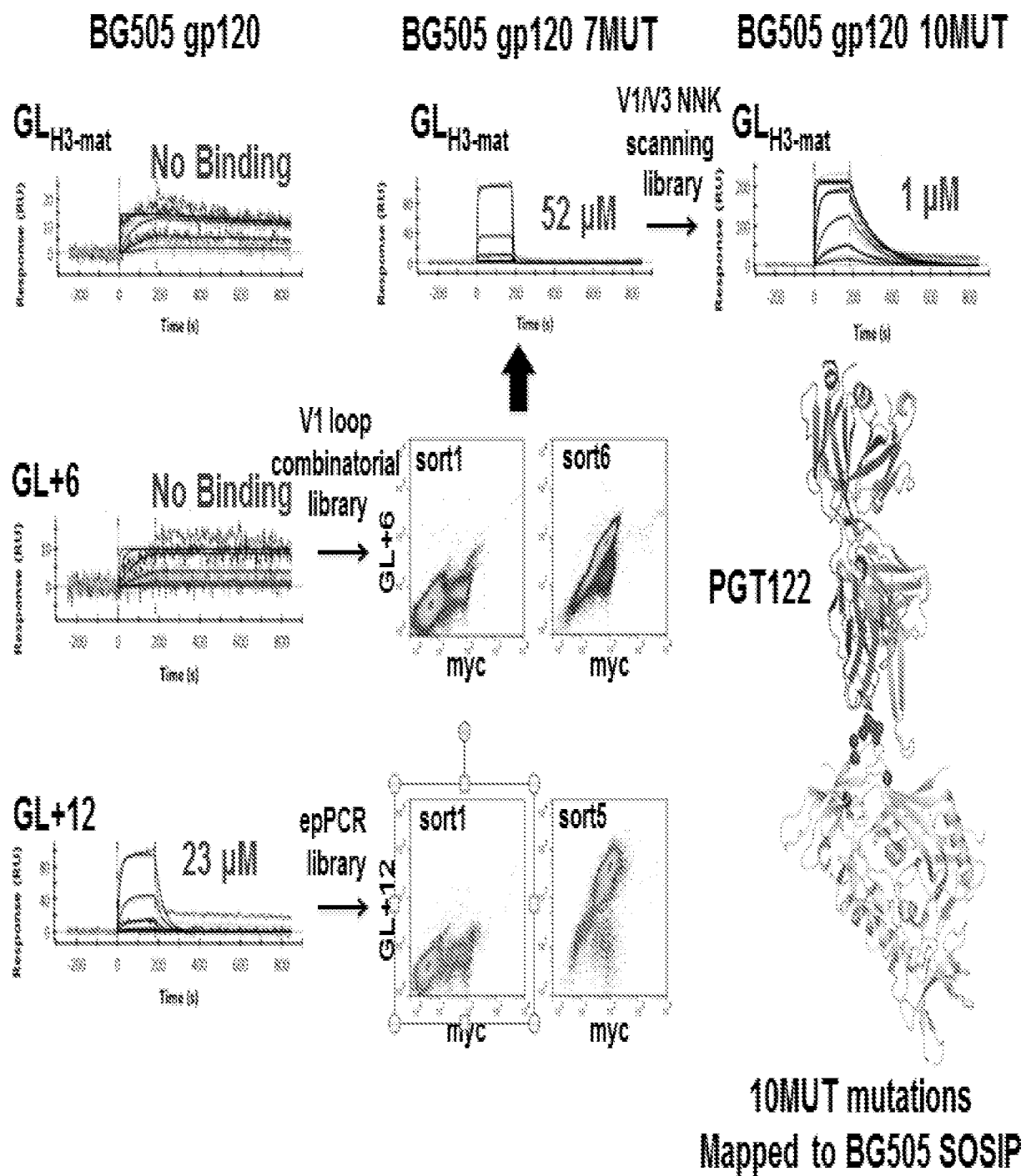


FIG. 2

BG505- gp120	PGT121	3H3L	GL+12+i ns	GL+12	GL+6	GL _{H3-mat}	GL _{H3-rev}	gH _{H3-rev} matL	mathH gL _{L3-rev}
WT	7	271	1500	~23000	NB	NB	NB	595	NB
3mut	4	9	229	2100	NB	NB	NB	22	NB
5mut	6	7	3	10	~208000		NB	4	~12000
7mut	0.21		<0.1	3	6430	~52000	NB	11	NB
10mut	0.09	<0.1			1230	1070	NB**	<0.1	~47000

FIG. 3

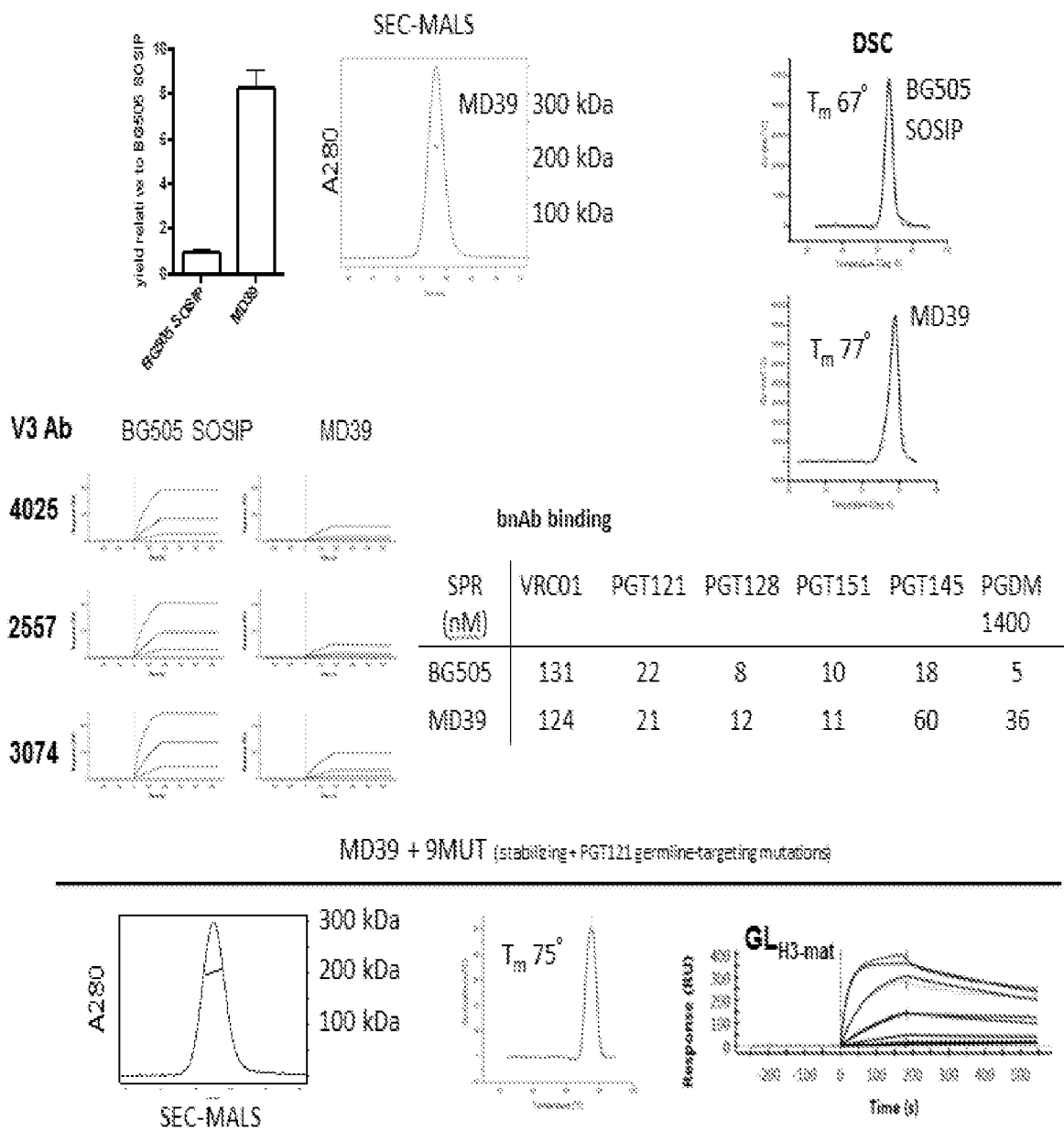


FIG. 4

5/42

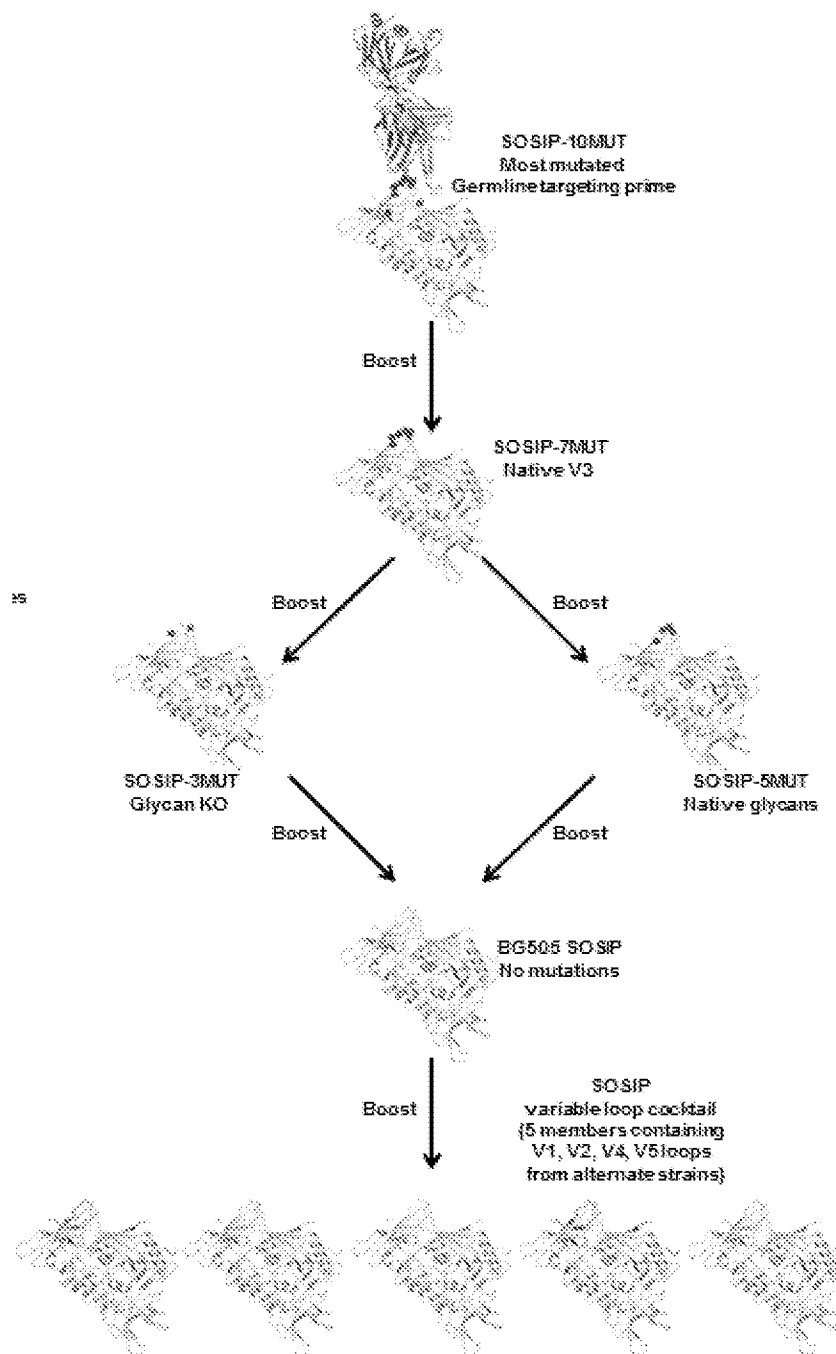
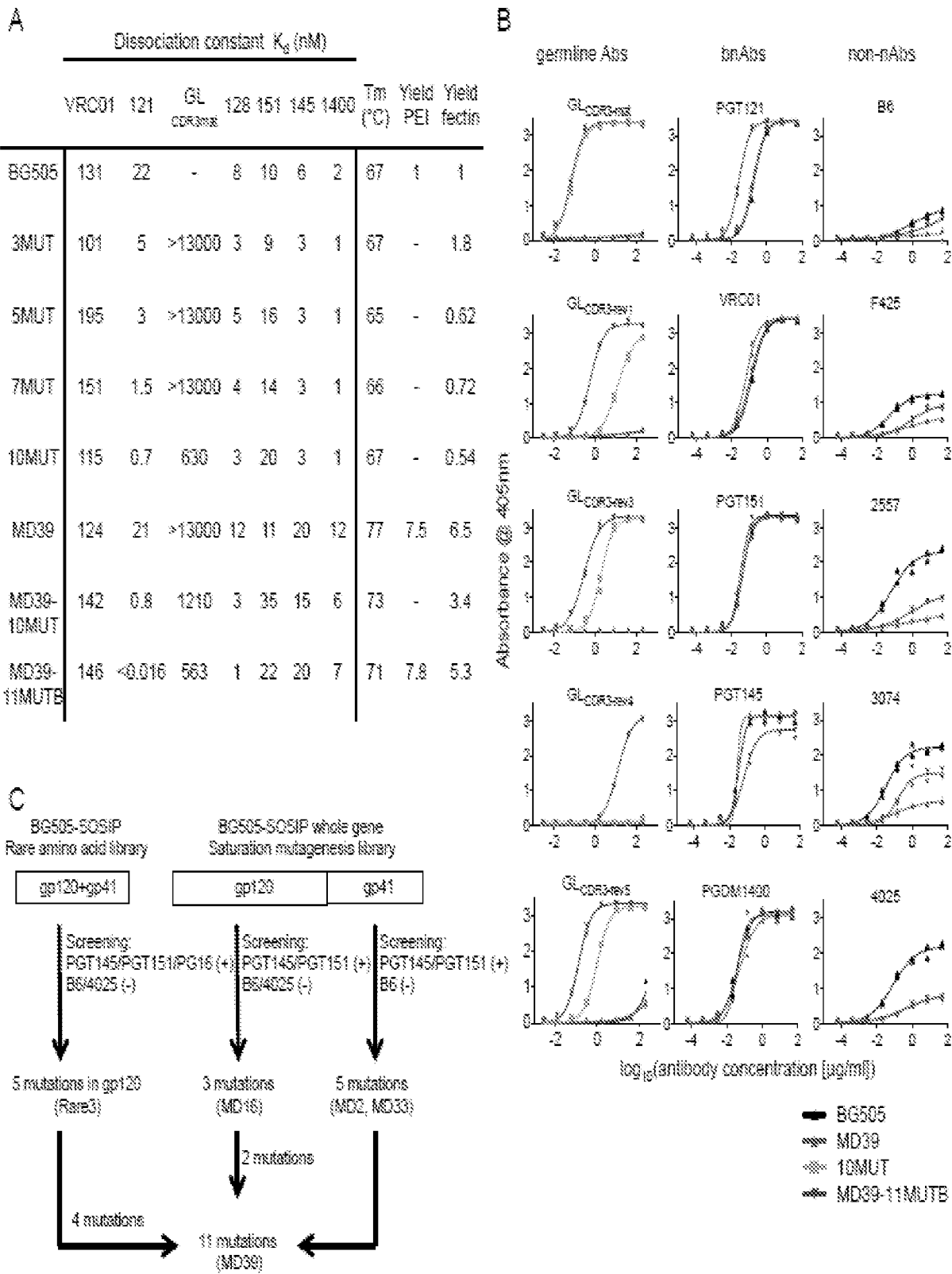
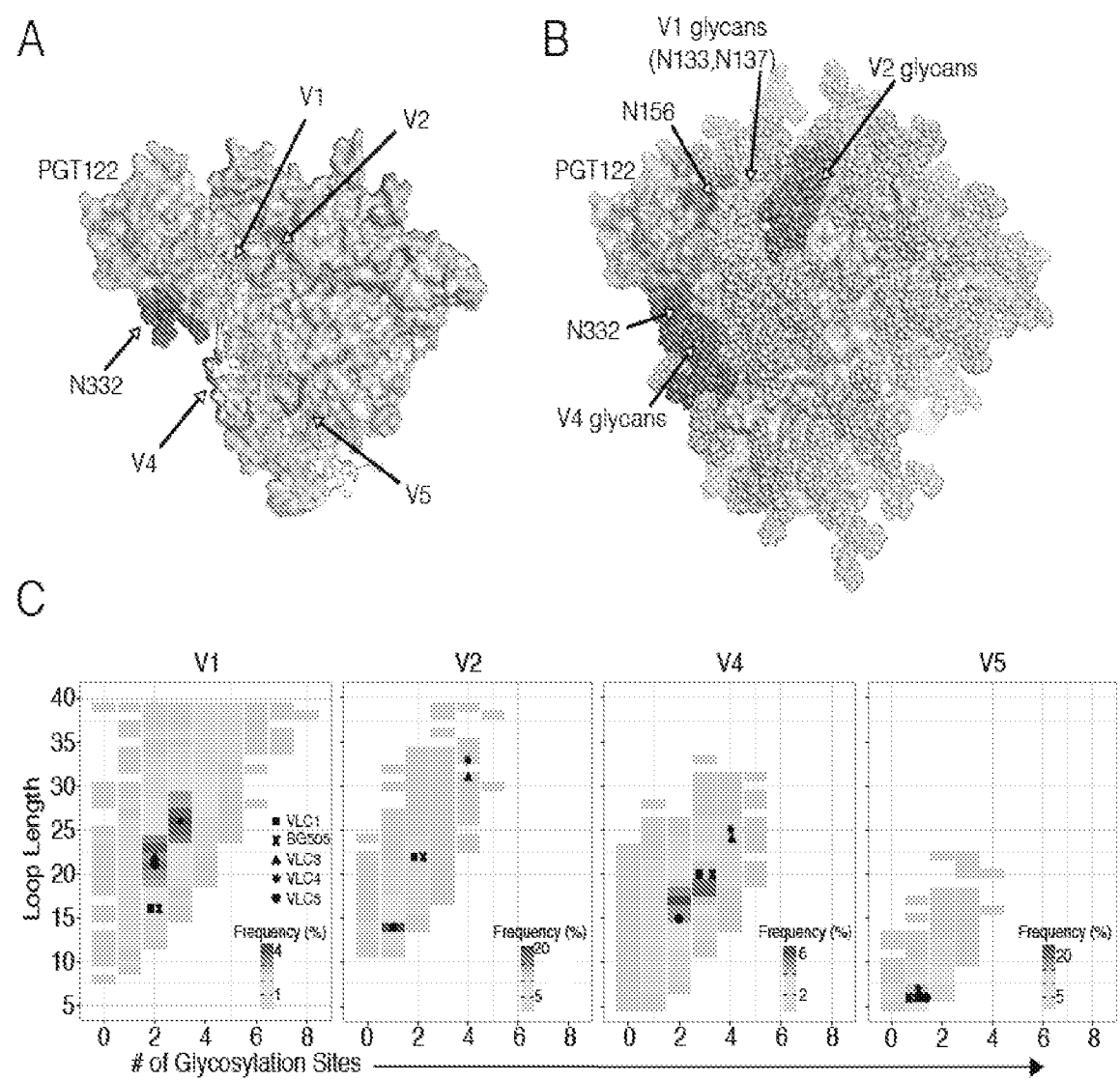


FIG. 5



FIGS. 7A-7C



FIGS. 8A-8C

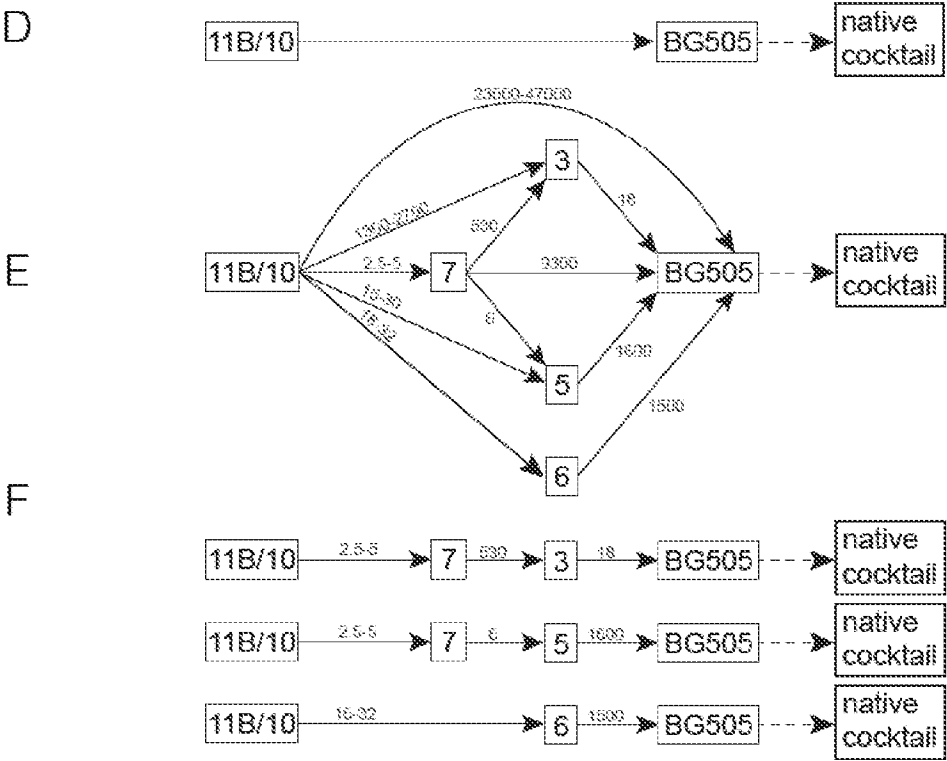


FIG. 8D-8F

10/42

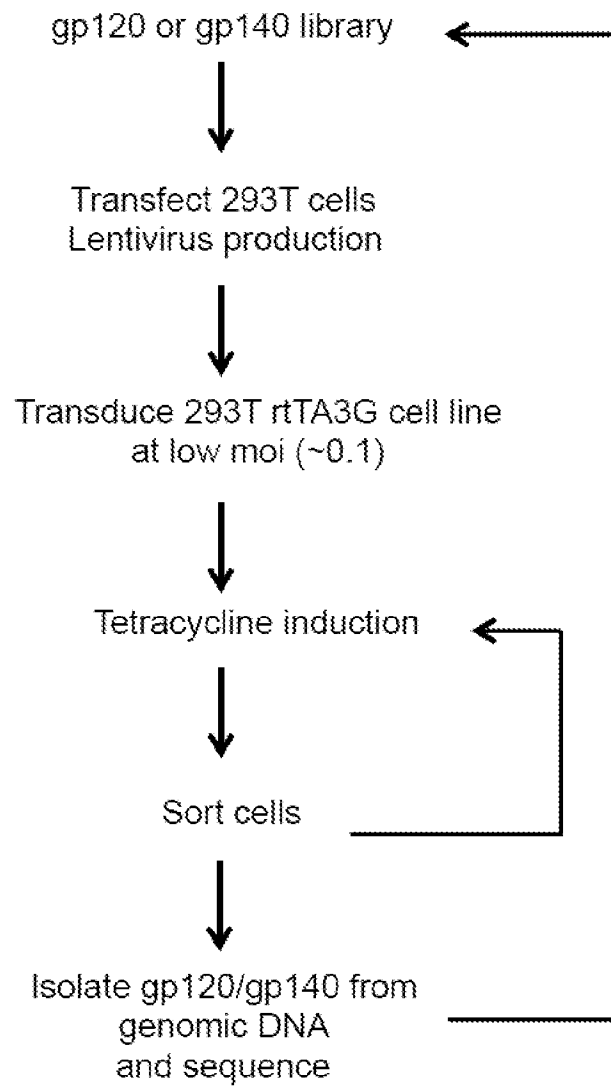


FIG. 9A

11/42

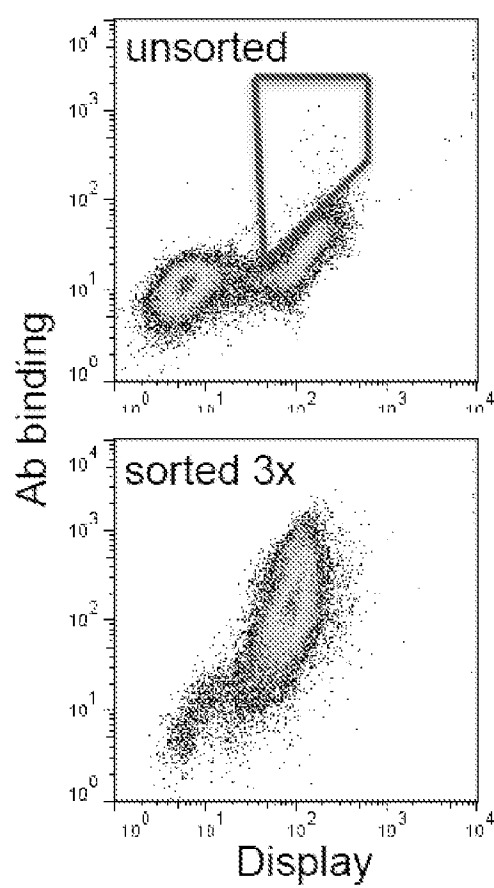


FIG. 9B

BG505-SOSIP.664_PDGER_TM

NDANNRGLOCVLLLOGAVFVSPPSQEIHARFRGAKAENLWVTVYVGVPVWKAETTLFCASDAKAYETEKHNWATHACVPTDPNPQEIHLENVTEEFN
 MNKNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELKNCSPNMTTELDRKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRL
 INCNTSAITQACPKVGFEPFIHYCAPAGFAILKCKDKKFNSTGCPSPSVSTVQCTHGINPVVSTQLLNGSLAESEVMIRSENITNNAKNILVQFNTFV
 QINCTRFNNNTKRSIRIGPGQAFYATGDIIGDIRQAHONVSKATWNETLGVVVKQLRKHFGNNTIIRFANSSSGGDLEVTTHSFNCGGEFFYCNSTGLFN
 STWISNTSVQGENSTGSNDSITLPCRIRQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNNTTETFRPGGGIMRDNRSELYKYVVKIE
 PLGVAPTRCKRNVVGGRRRRRAVGHGAVFLGFLCAAGSTWGAASMTLTVQARNLLSGIVQQSNLLRAPEAQGHLLKLTVMGKQLQARVLAVERYLRD
 QQLLGINGGSGKLIQCTNVFNWSSWENRNLSEINDMTWLQWDEEISNYTQIIYGLLEESQWQEKNEQDILLALDGGSGGGSGGSGQLISRRDLGGSG
 GSGGSNAVGQUTQEVIVVPHSLPFKVVVLSAILALVVLTIISLITLIMLWQKRP

BG505-gp120_PDGER_TM

MEITITLLLVLLNWPGSTGAENLWVTVYVGVPVWKAETTLFCASDAKAYETEKHNWATHACVPTDPNPQEIHLENVTEEFNMKNMVEQMHTD
 IISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELKNCSPNMTTELDRKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQAC
 PNVGFEPFIHYCAPAGFAILKCKDKKFNSTGCPSPSVSTVQCTHGINPVVSTQLLNGSLAESEVMIRSENITNNAKNILVQFNTFVQINCTRFNNNT
 RKSIRIGPGQAFYATGDIIGDIRQAHONVSKATWNETLGVVVKQLRKHFGNNTIIRFANSSSGGDLEVTTHSFNCGGEFFYCNSTGLFNSTWISNTSVQ
 GENSTGSNDSITLPCRIRQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNNTTETFRPGGGIMRDNRSELYKYVVKIEPLGVAPTRA
 KRRVVGSGSGSGSGGQLISRRDLGGSGGGGSNAVGQUTQEVIVVPHSLPFKVVVLSAILALVVLTIISLITLIMLWQKRP

FIG. 9C

13/42

Heavy Chain

V

V4-59/D3-3/J6	QVQLQESGPGLVKPSSETLSLTCTVSGGSISSYYWSWIRQPPGKGLEWIGYIYYSGSTNYPNPSLKSRVTIS
GL-CDR3rev1	QVQLQESGPGLVKPSSETLSLTCTVSGGSISSYYWSWIRQPPGKGLEWIGYIYYSGSTNYPNPSLKSRVTIS
GL-CDR3rev2	QVQLQESGPGLVKPSSETLSLTCTVSGGSISSYYWSWIRQPPGKGLEWIGYIYYSGSTNYPNPSLKSRVTIS
GL-CDR3rev3	QVQLQESGPGLVKPSSETLSLTCTVSGGSISSYYWSWIRQPPGKGLEWIGYIYYSGSTNYPNPSLKSRVTIS
GL-CDR3rev4	QVQLQESGPGLVKPSSETLSLTCTVSGGSISSYYWSWIRQPPGKGLEWIGYIYYSGSTNYPNPSLKSRVTIS
GL-CDR3rev5	QVQLQESGPGLVKPSSETLSLTCTVSGGSISSYYWSWIRQPPGKGLEWIGYIYYSGSTNYPNPSLKSRVTIS
GL+3	QVQLQESGPGLVKPSSETLSLTCTVSGGSISSYYWSWIRQPPGKGLEWIGYIYYSGSTNYPNPSLKSRVTIS
GL-CDR3mat	QVQLQESGPGLVKPSSETLSLTCTVSGGSISSYYWSWIRQPPGKGLEWIGYIYYSGSTNYPNPSLKSRVTIS
GL+9	QVQLQESGPGLVKPSSETLSLTCTVSGGSISSYYWSWIRQPPGKGLEWIGYIYYSGSTNYPNPSLKSRVTIS
PGT121	QVQLQESGPGLVKPSSETLSLTCTVSGGSISSYYWSWIRQPPGKGLEWIGYIYYSGSTNYPNPSLKSRVTIS

D

J

V4-59/D3-3/J6	VDTSKNQFSLKLSSVTAADTAVYYCAR----ITIFGVVIL--YYYYYYMDVWGKTTTIVTSS
GL-CDR3rev1	VDTSKNQFSLKLSSVTAADTAVYYCARITIGITIFGVVAFREYYYYYYMDVWGKTTTIVTSS
GL-CDR3rev2	VDTSKNQFSLKLSSVTAADTAVYYCARITIGITIFGVVILKEYYYYYYMDVWGKTTTIVTSS
GL-CDR3rev3	VDTSKNQFSLKLSSVTAADTAVYYCARITIGITIGVVSDEYYYYYYMDVWGKTTTIVTSS
GL-CDR3rev4	VDTSKNQFSLKLSSVTAADTAVYYCARITIGITIGVVSDEYYYYYYMDVWGKTTTIVTSS
GL-CDR3rev5	VDTSKNQFSLKLSSVTAADTAVYYCARITIGITIFGVVAFREYYYYYYMDVWGKTTTIVTSS
GL+3	VDTSKNQFSLKLSSVTAADTAVYYCARITIGITIGVVSDEYYYYYYMDVWGKTTTIVTSS
GL-CDR3mat	VDTSKNQFSLKLSSVTAADTAVYYCARITIGITIGVVSDEYYYYYYMDVWGKTTTIVTSS
GL+9	VDTSKNQFSLKLSSVTAADTAVYYCARITIGITIGVVSDEYYYYYYMDVWGKTTTIVTSS
PGT121	VDTSKNQFSLKLSSVTAADTAVYYCARITIGITIGVVSDEYYYYYYMDVWGKTTTIVTSS

Light Chain

V

V3-21/J3	SYVLTPPPSVSVAPGQTARITCGGNNIGSKSVHWYQQKPGQAPVLLVYDDSDRPSGIPERFSGSNS---G
GL-CDR3rev1	SYVLTPPPSVSVAPGQTARITCGGNNIGSKSVHWYQQKPGQAPVLLVYDDSDRPSGIPERFSGSNS---G
GL-CDR3rev2	SYVLTPPPSVSVAPGQTARITCGGNNIGSKSVHWYQQKPGQAPVLLVYDDSDRPSGIPERFSGSNS---G
GL-CDR3rev3	SYVLTPPPSVSVAPGQTARITCGGNNIGSKSVHWYQQKPGQAPVLLVYDDSDRPSGIPERFSGSNS---G
GL-CDR3rev4	SYVLTPPPSVSVAPGQTARITCGGNNIGSKSVHWYQQKPGQAPVLLVYDDSDRPSGIPERFSGSNS---G
GL-CDR3rev5	SYVLTPPPSVSVAPGQTARITCGGNNIGSKSVHWYQQKPGQAPVLLVYDDSDRPSGIPERFSGSNS---G
GL+3	SYVLTPPPSVSVAPGQTARITCGGNNIGSKSVHWYQQKPGQAPVLLVYDDSDRPSGIPERFSGSNS---G
GL-CDR3mat	SYVLTPPPSVSVAPGQTARITCGGNNIGSKSVHWYQQKPGQAPVLLVYDDSDRPSGIPERFSGSNS---G
GL+9	SYVLTPPPSVSVAPGQTARITCGGNNIGSKSVHWYQQKPGQAPVLLVYDDSDRPSGIPERFSGSNS---G
PGT121	-----SDISVAPGQTARITCGGNNIGSKSVHWYQQKPGQAPVLLVYDDSDRPSGIPERFSGSNS---G

J

V3-21/J3	NTATLTISRVEAGDEADYYCQVWDSSSDHPWVFGGGTKLTVL
GL-CDR3rev1	NTATLTISRVEAGDEADYYCQVWDSSSDHPWVFGGGTKLTVL
GL-CDR3rev2	NTATLTISRVEAGDEADYYCQVWDSSSDHPWVFGGGTKLTVL
GL-CDR3rev3	NTATLTISRVEAGDEADYYCQVWDSSSDHPWVFGGGTKLTVL
GL-CDR3rev4	NTATLTISRVEAGDEADYYCQVWDSSSDHPWVFGGGTKLTVL
GL-CDR3rev5	NTATLTISRVEAGDEADYYCQVWDSSSDHPWVFGGGTKLTVL
GL+3	NTATLTISRVEAGDEADYYCQVWDSSSDHPWVFGGGTKLTVL
GL-CDR3mat	NTATLTISRVEAGDEADYYCQVWDSSSDHPWVFGGGTKLTVL
GL+9	NTATLTISRVEAGDEADYYCQVWDSSSDHPWVFGGGTKLTVL
PGT121	NTATLTISRVEAGDEADYYCQVWDSSSDHPWVFGGGTKLTVL

FIG. 10

```
B5655 80S1P.664      IPIHYCAPAGFALLKCKDKFNKGTCPCFVSVTQCTHGIKFVSVTQLLNGSLABEEVMIRSEINTNNAKMILVQWNTVPQINCTRPNMNT
B5655 80S1P.664 2MTU IPIHYCAPAGFALLKCKDKFNKGTCPCFVSVTQCTHGIKFVSVTQLLNGSLABEEVMIRSEINTNNAKMILVQWNTVPQINCTRPNMNT
B5655 80S1P.664 3MTU IPIHYCAPAGFALLKCKDKFNKGTCPCFVSVTQCTHGIKFVSVTQLLNGSLABEEVMIRSEINTNNAKMILVQWNTVPQINCTRPNMNT
B5655 80S1P.664 5MTU IPIHYCAPAGFALLKCKDKFNKGTCPCFVSVTQCTHGIKFVSVTQLLNGSLABEEVMIRSEINTNNAKMILVQWNTVPQINCTRPNMNT
B5655 80S1P.664 6MTU IPIHYCAPAGFALLKCKDKFNKGTCPCFVSVTQCTHGIKFVSVTQLLNGSLABEEVMIRSEINTNNAKMILVQWNTVPQINCTRPNMNT
B5655 80S1P.664 7MTU IPIHYCAPAGFALLKCKDKFNKGTCPCFVSVTQCTHGIKFVSVTQLLNGSLABEEVMIRSEINTNNAKMILVQWNTVPQINCTRPNMNT
B5655 80S1P.664 9MTA IPIHYCAPAGFALLKCKDKFNKGTCPCFVSVTQCTHGIKFVSVTQLLNGSLABEEVMIRSEINTNNAKMILVQWNTVPQINCTRPNMNT
B5655 gp120_9MTU     IPIHYCAPAGFALLKCKDKFNKGTCPCFVSVTQCTHGIKFVSVTQLLNGSLABEEVMIRSEINTNNAKMILVQWNTVPQINCTRPNMNT
B5655 gp120_NKST_KO IPIHYCAPAGFALLKCKDKFNKGTCPCFVSVTQCTHGIKFVSVTQLLNGSLABEEVMIRSEINTNNAKMILVQWNTVPQINCTRPNMNT
B5655 gp120_1NMUT    IPIHYCAPAGFALLKCKDKFNKGTCPCFVSVTQCTHGIKFVSVTQLLNGSLABEEVMIRSEINTNNAKMILVQWNTVPQINCTRPNMNT
B5655 80S1P.664 16MTU IPIHYCAPAGFALLKCKDKFNKGTCPCFVSVTQCTHGIKFVSVTQLLNGSLABEEVMIRSEINTNNAKMILVQWNTVPQINCTRPNMNT
B5655 80S1P_MD39_10MTA IPIHYCAPAGFALLKCKDKFNKGTCPCFVSVTQCTHGIKFVSVTQLLNGSLABEEVMIRSEINTNNAKMILVQWNTVPQINCTRPNMNT
B5655 80S1P.664 18MTA IPIHYCAPAGFALLKCKDKFNKGTCPCFVSVTQCTHGIKFVSVTQLLNGSLABEEVMIRSEINTNNAKMILVQWNTVPQINCTRPNMNT
B5655 80S1P_MD39_14MTB IPIHYCAPAGFALLKCKDKFNKGTCPCFVSVTQCTHGIKFVSVTQLLNGSLABEEVMIRSEINTNNAKMILVQWNTVPQINCTRPNMNT
```

B5565 50S1P.464 RKSLIGPQQAIFYATGDIIGIDIRQAHCIVSKATWNETLGRVVKQLRKHFGNNTIIRFANSSGGGLVTTSHFNCGGEFFYCNISGLFNSTW
 B5565 50S1P.464 2MUT FSRIRIGQQAIFYATGDIIGIDIRQAHCIVSKATWNETLGRVVKQLRKHFGNNTIIRFANSSGGGLVTTSHFNCGGEFFYCNISGLFNSTW
 B5565 50S1P.464 3MUT RKRIRIGPQQAIFYATGDIIGIDIRQAHCIVSKATWNETLGRVVKQLRKHFGNNTIIRFANSSGGGLVTTSHFNCGGEFFYCNISGLFNSTW
 B5565 50S1P.464 5MUT RKRIRIGPQQAIFYATGDIIGIDIRQAHCIVSKATWNETLGRVVKQLRKHFGNNTIIRFANSSGGGLVTTSHFNCGGEFFYCNISGLFNSTW
 B5565 50S1P.464 7MUT RKRIRIGPQQAIFYATGDIIGIDIRQAHCIVSKATWNETLGRVVKQLRKHFGNNTIIRFANSSGGGLVTTSHFNCGGEFFYCNISGLFNSTW
 B5565 50S1P.464 10MUT RKRIRIGPQQAIFYATGDIIGIDIRQAHCIVSKATWNETLGRVVKQLRKHFGNNTIIRFANSSGGGLVTTSHFNCGGEFFYCNISGLFNSTW
 B5565 gp122 3MUTB RKRIRIGPQQAIFYATGDIIGIDIRQAHCIVSKATWNETLGRVVKQLRKHFGNNTIIRFANSSGGGLVTTSHFNCGGEFFYCNISGLFNSTW
 B5565 gp124 10MUT NC RKRIRIGPQQAIFYATGDIIGIDIRQAHCIVSKATWNETLGRVVKQLRKHFGNNTIIRFANSSGGGLVTTSHFNCGGEFFYCNISGLFNSTW
 B5565 gp122 3MUT RKRIRIGPQQAIFYATGDIIGIDIRQAHCIVSKATWNETLGRVVKQLRKHFGNNTIIRFANSSGGGLVTTSHFNCGGEFFYCNISGLFNSTW
 B5565 50S1P.464 10MUT FSRIRIGQQAIFYATGDIIGIDIRQAHCIVSKATWNETLGRVVKQLRKHFGNNTIIRFANSSGGGLVTTSHFNCGGEFFYCNISGLFNSTW
 B5565 50S1P_MB39 10MUTB RKRIRIGPQQAIFYATGDIIGIDIRQAHCIVSKATWNETLGRVVKQLRKHFGNNTIIRFANSSGGGLVTTSHFNCGGEFFYCNISGLFNSTW
 B5565 50S1P.464 10MUT FSRIRIGQQAIFYATGDIIGIDIRQAHCIVSKATWNETLGRVVKQLRKHFGNNTIIRFANSSGGGLVTTSHFNCGGEFFYCNISGLFNSTW
 B5565 50S1P_MB39 11MUTB RKRIRIGPQQAIFYATGDIIGIDIRQAHCIVSKATWNETLGRVVKQLRKHFGNNTIIRFANSSGGGLVTTSHFNCGGEFFYCNISGLFNSTW

```

B5565 _SOSIP..664      ISNTFVQGSNSTGSGNSITLPCRIFQILNNWQRIGQANYAPPICQVIRCVSNITGLILTRDGGSTNSTETTFPFGGGMRDNNRWSLEYLYXK
B5565 _SOSIP..664 2MTU  ISNTFVQGSNSTGSGNSITLPCRIFQILNNWQRIGQANYAPPICQVIRCVSNITGLILTRDGGSTNSTETTFPFGGGMRDNNRWSLEYLYXK
B5565 _SOSIP..664 3MTU  ISNTFVQGSNSTGSGNSITLPCRIFQILNNWQRIGQANYAPPICQVIRCVSNITGLILTRDGGSTNSTETTFPFGGGMRDNNRWSLEYLYXK
B5565 _SOSIP..664 5MTU  ISNTFVQGSNSTGSGNSITLPCRIFQILNNWQRIGQANYAPPICQVIRCVSNITGLILTRDGGSTNSTETTFPFGGGMRDNNRWSLEYLYXK
B5565 _SOSIP..664 6MTU  ISNTFVQGSNSTGSGNSITLPCRIFQILNNWQRIGQANYAPPICQVIRCVSNITGLILTRDGGSTNSTETTFPFGGGMRDNNRWSLEYLYXK
B5565 _SOSIP..664 7MTU  ISNTFVQGSNSTGSGNSITLPCRIFQILNNWQRIGQANYAPPICQVIRCVSNITGLILTRDGGSTNSTETTFPFGGGMRDNNRWSLEYLYXK
B5565 _SOSIP..664 8MTU  ISNTFVQGSNSTGSGNSITLPCRIFQILNNWQRIGQANYAPPICQVIRCVSNITGLILTRDGGSTNSTETTFPFGGGMRDNNRWSLEYLYXK
B5565 _SOSIP..664 9MTU  ISNTFVQGSNSTGSGNSITLPCRIFQILNNWQRIGQANYAPPICQVIRCVSNITGLILTRDGGSTNSTETTFPFGGGMRDNNRWSLEYLYXK
B5565 gp120 2MTU  ISNTFVQGSNSTGSGNSITLPCRIFQILNNWQRIGQANYAPPICQVIRCVSNITGLILTRDGGSTNSTETTFPFGGGMRDNNRWSLEYLYXK
B5565 gp120 10MTU  ISNTFVQGSNSTGSGNSITLPCRIFQILNNWQRIGQANYAPPICQVIRCVSNITGLILTRDGGSTNSTETTFPFGGGMRDNNRWSLEYLYXK
B5565 gp120 18MTU  ISNTFVQGSNSTGSGNSITLPCRIFQILNNWQRIGQANYAPPICQVIRCVSNITGLILTRDGGSTNSTETTFPFGGGMRDNNRWSLEYLYXK
B5565 _SOSIP..664 10MTU  ISNTFVQGSNSTGSGNSITLPCRIFQILNNWQRIGQANYAPPICQVIRCVSNITGLILTRDGGSTNSTETTFPFGGGMRDNNRWSLEYLYXK
B5565 _SOSIP..MD39 10MTU  ISNTFVQGSNSTGSGNSITLPCRIFQILNNWQRIGQANYAPPICQVIRCVSNITGLILTRDGGSTNSTETTFPFGGGMRDNNRWSLEYLYXK
B5565 _SOSIP..664 11MTU  ISNTFVQGSNSTGSGNSITLPCRIFQILNNWQRIGQANYAPPICQVIRCVSNITGLILTRDGGSTNSTETTFPFGGGMRDNNRWSLEYLYXK
B5565 _SOSIP..MD39 11MTU  ISNTFVQGSNSTGSGNSITLPCRIFQILNNWQRIGQANYAPPICQVIRCVSNITGLILTRDGGSTNSTETTFPFGGGMRDNNRWSLEYLYXK

```

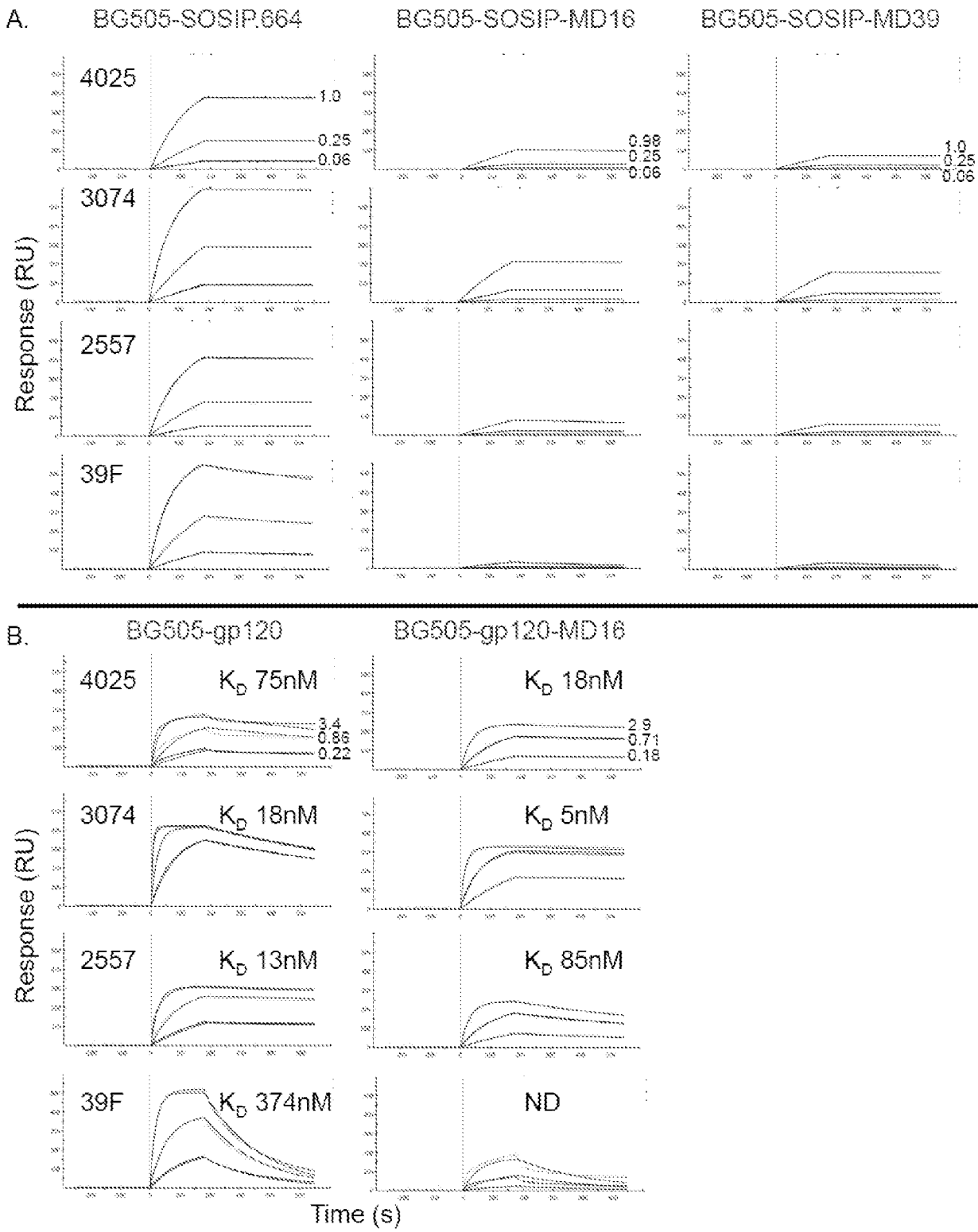
```

B5505 SOSIP.654 VVKIEPLGVAPTRCKRRVYVGRRRRPAVGIQAVFLGFLGAASTNGAASNTLTVAQRNLLSGIVQQQSNLLRAPEAQHLLKLTVWGIGQL
B5505 SOSIP.654 2MUT VKKIEPLGVAPTRCKRRVYVGRRRRPAVGIQAVFLGFLGAASTNGAASNTLTVAQRNLLSGIVQQQSNLLRAPEAQHLLKLTVWGIGQL
B5505 SOSIP.654 3MUT VVKIEPLGVAPTRCKRRVYVGRRRRPAVGIQAVFLGFLGAASTNGAASNTLTVAQRNLLSGIVQQQSNLLRAPEAQHLLKLTVWGIGQL
B5505 SOSIP.654 5MUT VVKIEPLGVAPTRCKRRVYVGRRRRPAVGIQAVFLGFLGAASTNGAASNTLTVAQRNLLSGIVQQQSNLLRAPEAQHLLKLTVWGIGQL
B5505 SOSIP.654 6MUT VVKIEPLGVAPTRCKRRVYVGRRRRPAVGIQAVFLGFLGAASTNGAASNTLTVAQRNLLSGIVQQQSNLLRAPEAQHLLKLTVWGIGQL
B5505 SOSIP.654 7MUT VVKIEPLGVAPTRCKRRVYVGRRRRPAVGIQAVFLGFLGAASTNGAASNTLTVAQRNLLSGIVQQQSNLLRAPEAQHLLKLTVWGIGQL
B5505 SOSIP.654 9MUT VVKIEPLGVAPTRCKRRVYVGRRRRPAVGIQAVFLGFLGAASTNGAASNTLTVAQRNLLSGIVQQQSNLLRAPEAQHLLKLTVWGIGQL
B5505 gp120 2MUTB VVKIEP
B5505 gp120 10MUT NO VVKIEP
B5505 gp120 13MUT VVKIEP
B5505 SOSIP.654 16MUT VVKIEPLGVAPTRCKRRVYVGRRRRPAVGIQAVFLGFLGAASTNGAASNTLTVAQRNLLSGIVQQQSNLLRAPEAQHLLKLTVWGIGQL
B5505 SOSIP.MD39 13MUTB VVKIEPLGVAPTRCKRRVYVGRRRRPAVGIQAVFLGFLGAASTNGAASNTLTVAQRNLLSGIVQQQSNLLRAPEAQHLLKLTVWGIGQL
B5505 SOSIP.654 14MUTB VVKIEPLGVAPTRCKRRVYVGRRRRPAVGIQAVFLGFLGAASTNGAASNTLTVAQRNLLSGIVQQQSNLLRAPEAQHLLKLTVWGIGQL
B5505 SOSIP.MD39 14MUTB VVKIEPLGVAPTRCKRRVYVGRRRRPAVGIQAVFLGFLGAASTNGAASNTLTVAQRNLLSGIVQQQSNLLRAPEAQHLLKLTVWGIGQL

```

[illegible]

FIG. 11



FIGS. 12A-12B

BG505_SOSIP_664 AENLWVTVYYGVFVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPPNFQEIHLNVTEEFNMWKNMVEQMHDTIIS
 BG505_SOSIP_rare3 AENLWVTVYYGVFVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPPNFQEIHLNVTEEFNMWKNMVEQMHDTIIS
 BG505_SOSIP_MD16 AENLWVTVYYGVFVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPPNFQEIHLNVTEEFNMWKNMVEQMHDTIIS
 BG505_SOSIP_MD2 AENLWVTVYYGVFVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPPNFQEIHLNVTEEFNMWKNMVEQMHDTIIS
 BG505_SOSIP_MD33 AENLWVTVYYGVFVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPPNFQEIHLNVTEEFNMWKNMVEQMHDTIIS
 BG505_SOSIP_MD39 AENLWVTVYYGVFVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPPNFQEIHLNVTEEFNMWKNMVEQMHDTIIS

BG505_SOSIP_664 LWDQSLKPCVKLTPLCVTLQCTNVTNNITDMDRGELKNCSPNMTTELADKKQKVYSLFYRLDVVQINENQGNRSNNSNKE
 BG505_SOSIP_rare3 LWDQSLKPCVKLTPLCVTLQCTNVTNNITDMDRGELKNCSPNMTTELADKKQKVYSLFYRLDVVQINENQGNRSNNSNKE
 BG505_SOSIP_MD16 LWDQSLKPCVKLTPLCVTLQCTNVTNNITDMDRGELKNCSPNMTTELADKKQKVYSLFYRLDVVQINENQGNRSNNSNKE
 BG505_SOSIP_MD2 LWDQSLKPCVKLTPLCVTLQCTNVTNNITDMDRGELKNCSPNMTTELADKKQKVYSLFYRLDVVQINENQGNRSNNSNKE
 BG505_SOSIP_MD33 LWDQSLKPCVKLTPLCVTLQCTNVTNNITDMDRGELKNCSPNMTTELADKKQKVYSLFYRLDVVQINENQGNRSNNSNKE
 BG505_SOSIP_MD39 LWDQSLKPCVKLTPLCVTLQCTNVTNNITDMDRGELKNCSPNMTTELADKKQKVYSLFYRLDVVQINENQGNRSNNSNKE

BG505_SOSIP_664 YRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDKKFNCTGFCPSVSVTVQCTHGKIPVVSQJLLNGSLAEDEV
 BG505_SOSIP_rare3 YRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDKKFNCTGFCPSVSVTVQCTHGKIPVVSQJLLNGSLAEDEV
 BG505_SOSIP_MD16 YRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDKKFNCTGFCPSVSVTVQCTHGKIPVVSQJLLNGSLAEDEV
 BG505_SOSIP_MD2 YRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDKKFNCTGFCPSVSVTVQCTHGKIPVVSQJLLNGSLAEDEV
 BG505_SOSIP_MD33 YRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDKKFNCTGFCPSVSVTVQCTHGKIPVVSQJLLNGSLAEDEV
 BG505_SOSIP_MD39 YRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDKKFNCTGFCPSVSVTVQCTHGKIPVVSQJLLNGSLAEDEV

BG505_SOSIP_664 MIRSENITNNAKNILVQFNTPVQINCTRPNNNTKRSIRIGPGQAFYATGDIIGDIRQAHNCVSKATWNETLGVVVKQLRK
 BG505_SOSIP_rare3 MIRSENITNNAKNILVQFNTPVQINCTRPNNNTKRSIRIGPGQAFYATGDIIGDIRQAHNCVSKATWNETLGVVVKQLRK
 BG505_SOSIP_MD16 MIRSENITNNAKNILVQFNTPVQINCTRPNNNTKRSIRIGPGQAFYATGDIIGDIRQAHNCVSKATWNETLGVVVKQLRK
 BG505_SOSIP_MD2 MIRSENITNNAKNILVQFNTPVQINCTRPNNNTKRSIRIGPGQAFYATGDIIGDIRQAHNCVSKATWNETLGVVVKQLRK
 BG505_SOSIP_MD33 MIRSENITNNAKNILVQFNTPVQINCTRPNNNTKRSIRIGPGQAFYATGDIIGDIRQAHNCVSKATWNETLGVVVKQLRK
 BG505_SOSIP_MD39 MIRSENITNNAKNILVQFNTPVQINCTRPNNNTKRSIRIGPGQAFYATGDIIGDIRQAHNCVSKATWNETLGVVVKQLRK

BG505_SOSIP_664 HFGNNTIIRFANSOGDLEVTTHSFNCGGEFFYCNSTGLFNSTWISNTSVQGSNSTGENDSITLPCRIKQIINMWQRIGQ
 BG505_SOSIP_rare3 HFGNNTIIRFANSOGDLEVTTHSFNCGGEFFYCNSTGLFNSTWISNTSVQGSNSTGENDSITLPCRIKQIINMWQRIGQ
 BG505_SOSIP_MD16 HFGNNTIIRFANSOGDLEVTTHSFNCGGEFFYCNSTGLFNSTWISNTSVQGSNSTGENDSITLPCRIKQIINMWQRIGQ
 BG505_SOSIP_MD2 HFGNNTIIRFANSOGDLEVTTHSFNCGGEFFYCNSTGLFNSTWISNTSVQGSNSTGENDSITLPCRIKQIINMWQRIGQ
 BG505_SOSIP_MD33 HFGNNTIIRFANSOGDLEVTTHSFNCGGEFFYCNSTGLFNSTWISNTSVQGSNSTGENDSITLPCRIKQIINMWQRIGQ
 BG505_SOSIP_MD39 HFGNNTIIRFANSOGDLEVTTHSFNCGGEFFYCNSTGLFNSTWISNTSVQGSNSTGENDSITLPCRIKQIINMWQRIGQ

BG505_SOSIP_664 AMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGGGMDRDNWRSELYKYKVVKIEPLGVAPTCKRRVVGRRRRR
 BG505_SOSIP_rare3 AMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGGGMDRDNWRSELYKYKVVKIEPLGVAPTCKRRVVGRRRRR
 BG505_SOSIP_MD16 AMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGGGMDRDNWRSELYKYKVVKIEPLGVAPTCKRRVVGRRRRR
 BG505_SOSIP_MD2 AMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGGGMDRDNWRSELYKYKVVKIEPLGVAPTCKRRVVGRRRRR
 BG505_SOSIP_MD33 AMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGGGMDRDNWRSELYKYKVVKIEPLGVAPTCKRRVVGRRRRR
 BG505_SOSIP_MD39 AMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGGGMDRDNWRSELYKYKVVKIEPLGVAPTCKRRVVGRRRRR

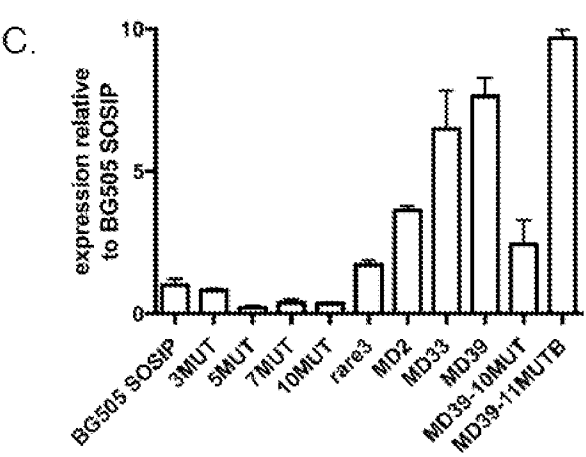
BG505_SOSIP_664 RAVGIGAVFLGFLGAAGSTMGAASMTLTQARNLLSGIVQQSNLLRAPEAQHLLKLTWVGIKQLQARVLAVERYLRDQ
 BG505_SOSIP_rare3 RAVGIGAVFLGFLGAAGSTMGAASMTLTQARNLLSGIVQQSNLLRAPEAQHLLKLTWVGIKQLQARVLAVERYLRDQ
 BG505_SOSIP_MD16 RAVGIGAVFLGFLGAAGSTMGAASMTLTQARNLLSGIVQQSNLLRAPEAQHLLKLTWVGIKQLQARVLAVERYLRDQ
 BG505_SOSIP_MD2 RAVGIGAVFLGFLGAAGSTMGAASMTLTQARNLLSGIVQQSNLLRAPEAQHLLKLTWVGIKQLQARVLAVERYLRDQ
 BG505_SOSIP_MD33 RAVGIGAVFLGFLGAAGSTMGAASMTLTQARNLLSGIVQQSNLLRAPEAQHLLKLTWVGIKQLQARVLAVERYLRDQ
 BG505_SOSIP_MD39 RAVGIGAVFLGFLGAAGSTMGAASMTLTQARNLLSGIVQQSNLLRAPEAQHLLKLTWVGIKQLQARVLAVERYLRDQ

BG505_SOSIP_664 QLLGIWGCSSGLICCTNVPWNSWSNRLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNGQEKNEQDILLALD
 BG505_SOSIP_rare3 QLLGIWGCSSGLICCTNVPWNSWSNRLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNGQEKNEQDILLALD
 BG505_SOSIP_MD16 QLLGIWGCSSGLICCTNVPWNSWSNRLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNGQEKNEQDILLALD
 BG505_SOSIP_MD2 QLLGIWGCSSGLICCTNVPWNSWSNRLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNGQEKNEQDILLALD
 BG505_SOSIP_MD33 QLLGIWGCSSGLICCTNVPWNSWSNRLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNGQEKNEQDILLALD
 BG505_SOSIP_MD39 QLLGIWGCSSGLICCTNVPWNSWSNRLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNGQEKNEQDILLALD

FIG. 13A

B.

Trimer	Tm (°C)
BG505 SOSIP	66.65
rare3	68.08
MD16	66.5
MD2	66.99
MD33	70.65
MD39	77.17



FIGS. 13B-13C

	Dissociation constant K_d (nM)						T _m (°C)	Yield PEI
	VRC01	121	128	151	145	1400		
BG505	131	22	8	10	6	2	66	1
MD39-VLC1	416	42	5	10	31	-	73	13.7
MD39-VLC2	124	21	12	11	20	12	77	6.2
MD39-VLC3	449	15	7	11	33	-	77	7.0
MD39-VLC4	167	13	2	11	27	-	76	6.3
MD39-VLC5	121	50	7	10	11	-	74	10.5

FIG. 14A

BG505_SOSIP_D664_MD39 AENLWVTVYVGVPVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQSIHLENVTEEFNMKNMNMVE
 BG505_SOSIP_MD39_VLC1-03 AENLWVTVYVGVPVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQSIHLENVTEEFNMKNMNMVE
 BG505_SOSIP_MD39_VLC2-04 AENLWVTVYVGVPVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQSIHLENVTEEFNMKNMNMVE
 BG505_SOSIP_MD39_VLC2-08 AENLWVTVYVGVPVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQSIHLENVTEEFNMKNMNMVE
 BG505_SOSIP_MD39_VLC3-13 AENLWVTVYVGVPVWKDAETTLFCASDAKAYETEKHNWATHACVPTDPNPQSIHLENVTEEFNMKNMNMVE

 BG505_SOSIP_D664_MD39 QMHEDIISLWDQSLNFCVKLTPLCVTLQC-----TNVTNNITDS-----MRGELKNCSPNNMTTELADKXQNV
 BG505_SOSIP_MD39_VLC1-03 QMHEDIISLWDQSLNFCVKLTPLCVTLQC-----TNVDNATLAN-----MRGELKNCSPNNMTTELADKXQNV
 BG505_SOSIP_MD39_VLC2-04 QMHEDIISLWDQSLNFCVKLTPLCVTLQC-----TNVDNATLAN-----MRGELKNCSPNNMTTELADKXQNV
 BG505_SOSIP_MD39_VLC2-08 QMHEDIISLWDQSLNFCVKLTPLCVTLQC-----TNVDNATLAN-----MRGELKNCSPNNMTTELADKXQNV
 BG505_SOSIP_MD39_VLC3-13 QMHEDIISLWDQSLNFCVKLTPLCVTLQC-----TNVDNATLAN-----MRGELKNCSPNNMTTELADKXQNV

 BG505_SOSIP_D664_MD39 YSLFYRLDVVOINEN-----QGNKSNNSNKEYRLINCNTSAITQACPXVSFEFPIPIHYCAPAGFAI
 BG505_SOSIP_MD39_VLC1-03 YSLFYRLDVVOINEN-----QGNKSNNSNKEYRLINCNTSAITQACPXVSFEFPIPIHYCAPAGFAI
 BG505_SOSIP_MD39_VLC2-04 YSLFYRLDVVOINEN-----QGNKSNNSNKEYRLINCNTSAITQACPXVSFEFPIPIHYCAPAGFAI
 BG505_SOSIP_MD39_VLC2-08 YSLFYRLDVVOINEN-----QGNKSNNSNKEYRLINCNTSAITQACPXVSFEFPIPIHYCAPAGFAI
 BG505_SOSIP_MD39_VLC3-13 YSLFYRLDVVOINEN-----QGNKSNNSNKEYRLINCNTSAITQACPXVSFEFPIPIHYCAPAGFAI

 BG505_SOSIP_D664_MD39 LKCKDKFNFGTGPCPSVSTVQCTHGIKFPVSTQLLNGSLABEEVIRSENITNNAKNILVQLNTFVQINCT
 BG505_SOSIP_MD39_VLC1-03 LKCKDKFNFGTGPCPSVSTVQCTHGIKFPVSTQLLNGSLABEEVIRSENITNNAKNILVQLNTFVQINCT
 BG505_SOSIP_MD39_VLC2-04 LKCKDKFNFGTGPCPSVSTVQCTHGIKFPVSTQLLNGSLABEEVIRSENITNNAKNILVQLNTFVQINCT
 BG505_SOSIP_MD39_VLC2-08 LKCKDKFNFGTGPCPSVSTVQCTHGIKFPVSTQLLNGSLABEEVIRSENITNNAKNILVQLNTFVQINCT
 BG505_SOSIP_MD39_VLC3-13 LKCKDKFNFGTGPCPSVSTVQCTHGIKFPVSTQLLNGSLABEEVIRSENITNNAKNILVQLNTFVQINCT

 BG505_SOSIP_D664_MD39 RPNNTVKSIRIGPGQAFYTTGDIIGDIRQAHNCVSKATWNETLGVVVKQLRKHFGNNTIIRFAQSSGGDL
 BG505_SOSIP_MD39_VLC1-03 RPNNTVKSIRIGPGQAFYTTGDIIGDIRQAHNCVSKATWNETLGVVVKQLRKHFGNNTIIRFAQSSGGDL
 BG505_SOSIP_MD39_VLC2-04 RPNNTVKSIRIGPGQAFYTTGDIIGDIRQAHNCVSKATWNETLGVVVKQLRKHFGNNTIIRFAQSSGGDL
 BG505_SOSIP_MD39_VLC2-08 RPNNTVKSIRIGPGQAFYTTGDIIGDIRQAHNCVSKATWNETLGVVVKQLRKHFGNNTIIRFAQSSGGDL
 BG505_SOSIP_MD39_VLC3-13 RPNNTVKSIRIGPGQAFYTTGDIIGDIRQAHNCVSKATWNETLGVVVKQLRKHFGNNTIIRFAQSSGGDL

 BG505_SOSIP_D664_MD39 VTTNFPNCGEFFYCNSTGLFNSTWISNT-----SVQSGNSTGSDSITLPCRIKQIIMMWQRIQANYAPP
 BG505_SOSIP_MD39_VLC1-03 VTTNFPNCGEFFYCNSTGLFNSTWISNT-----SVQSGNSTGSDSITLPCRIKQIIMMWQRIQANYAPP
 BG505_SOSIP_MD39_VLC2-04 VTTNFPNCGEFFYCNSTGLFNSTWISNT-----SVQSGNSTGSDSITLPCRIKQIIMMWQRIQANYAPP
 BG505_SOSIP_MD39_VLC2-08 VTTNFPNCGEFFYCNSTGLFNSTWISNT-----SVQSGNSTGSDSITLPCRIKQIIMMWQRIQANYAPP
 BG505_SOSIP_MD39_VLC3-13 VTTNFPNCGEFFYCNSTGLFNSTWISNT-----SVQSGNSTGSDSITLPCRIKQIIMMWQRIQANYAPP

 BG505_SOSIP_D664_MD39 IQGVIRCVSNITGLILTRDGGSTN-STTETFRPGGGDMRDNRSELKYKVVVKIEPLGVAPTRCKRRVVGR
 BG505_SOSIP_MD39_VLC1-03 IQGVIRCVSNITGLILTRDGGSTN-STTETFRPGGGDMRDNRSELKYKVVVKIEPLGVAPTRCKRRVVGR
 BG505_SOSIP_MD39_VLC2-04 IQGVIRCVSNITGLILTRDGGSTN-STTETFRPGGGDMRDNRSELKYKVVVKIEPLGVAPTRCKRRVVGR
 BG505_SOSIP_MD39_VLC2-08 IQGVIRCVSNITGLILTRDGGSTN-STTETFRPGGGDMRDNRSELKYKVVVKIEPLGVAPTRCKRRVVGR
 BG505_SOSIP_MD39_VLC3-13 IQGVIRCVSNITGLILTRDGGSTN-STTETFRPGGGDMRDNRSELKYKVVVKIEPLGVAPTRCKRRVVGR

 BG505_SOSIP_D664_MD39 RRRRAVGIGAVSLGFLGAAGSTMGAAASNTLTVDARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGINKQLAR
 BG505_SOSIP_MD39_VLC1-03 RRRRAVGIGAVSLGFLGAAGSTMGAAASNTLTVDARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGINKQLAR
 BG505_SOSIP_MD39_VLC2-04 RRRRAVGIGAVSLGFLGAAGSTMGAAASNTLTVDARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGINKQLAR
 BG505_SOSIP_MD39_VLC2-08 RRRRAVGIGAVSLGFLGAAGSTMGAAASNTLTVDARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGINKQLAR
 BG505_SOSIP_MD39_VLC3-13 RRRRAVGIGAVSLGFLGAAGSTMGAAASNTLTVDARNLLSGIVQQQSNLLRAPEPQQHLLKDTHWGINKQLAR

 BG505_SOSIP_D664_MD39 VLAVERHYLRDQQLLGIWGCSEKLI CCTNVFWNSWSNKNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQN
 BG505_SOSIP_MD39_VLC1-03 VLAVERHYLRDQQLLGIWGCSEKLI CCTNVFWNSWSNKNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQN
 BG505_SOSIP_MD39_VLC2-04 VLAVERHYLRDQQLLGIWGCSEKLI CCTNVFWNSWSNKNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQN
 BG505_SOSIP_MD39_VLC2-08 VLAVERHYLRDQQLLGIWGCSEKLI CCTNVFWNSWSNKNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQN
 BG505_SOSIP_MD39_VLC3-13 VLAVERHYLRDQQLLGIWGCSEKLI CCTNVFWNSWSNKNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQN

 BG505_SOSIP_D664_MD39 QQEKNEQDLALD
 BG505_SOSIP_MD39_VLC1-03 QQEKNEQDLALD
 BG505_SOSIP_MD39_VLC2-04 QQEKNEQDLALD
 BG505_SOSIP_MD39_VLC2-08 QQEKNEQDLALD
 BG505_SOSIP_MD39_VLC3-13 QQEKNEQDLALD

FIG. 14B

20/42

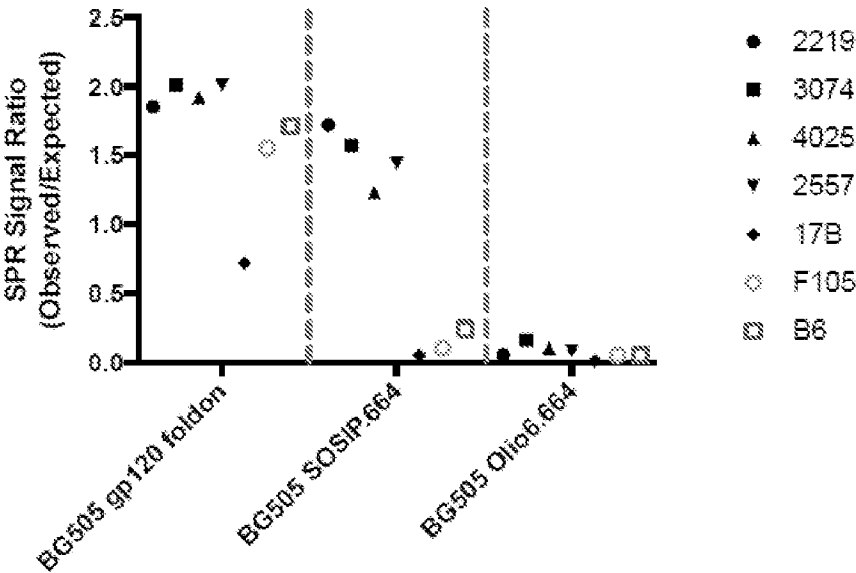
Position	N332-epitope Diversity
156	N(97%)
158	S(89%), T(10%)
295	N(88%), V(13%)
297	T(81%), I(12%)
299	P(96%)
301	N(95%)
303	T(94%)
320	T(90%)
321	G(79%)
322	D(49%), E(22%)
323	I(85%)
324	G(96%)
325	D(80%), N(16%)
326	I(96%)
327	R(98%)
328	Q(65%), K(24%)
330	H(65%), Y(31%)
332	N(72%), E(12%)
334	S(67%), N(29%)
386	N(90%)
388	T(48%), S(48%)
389	Q(26%), K(25%), G(19%)
392	N(92%)
394	T(79%)
415	T(71%), I(22%)
417	P(79%), Q(16%)
419	R(81%), K(18%)
442	N(27%), Q(22%)
444	T(29%), R(20%), N(20%)

FIG. 14C

A. Binding to bnAbs assessed by SPR

KD (nM)	V1V2 apex		N332 supersite		CD4 binding site		gp120-gp41 interface
	PGDM 1400	PGT 145	PGT 121	PGT 128	VRC01	3BNC60	PGT 151
BG505 SOSIP.664	6	15	26	8	131	137	10
BG505 MD39	12	20	21	11	124	189	11
BG505 Olig6	19	31	15	8	90	119	16

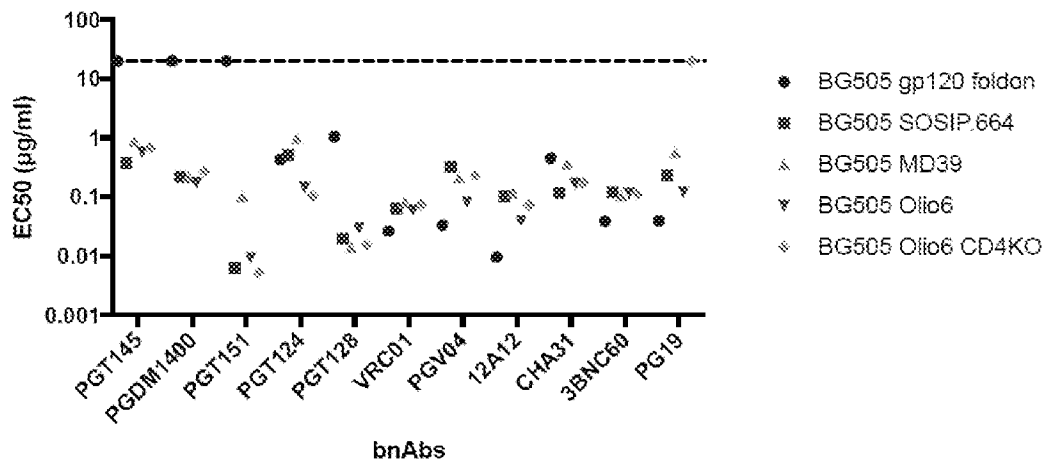
B. Binding to non-nAbs assessed by SPR



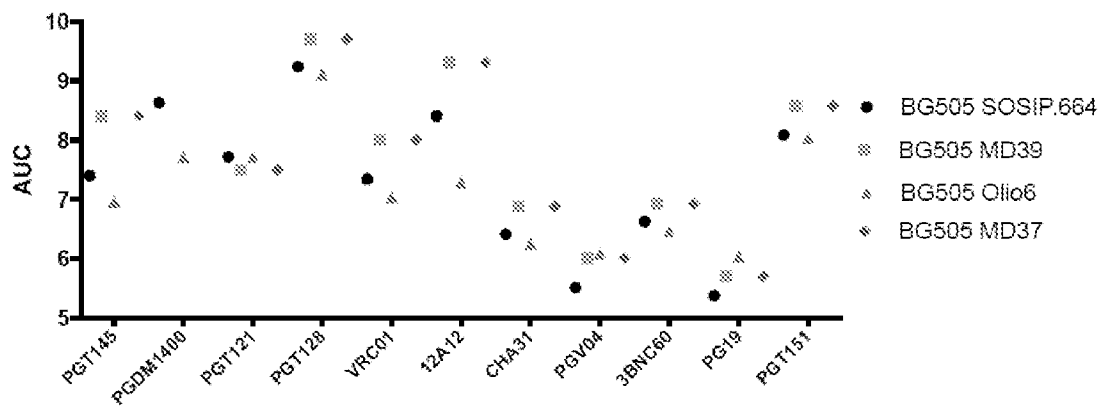
FIGS. 15A-15B

22/42

C. Binding to bnAbs assessed by ELISA, data as EC50

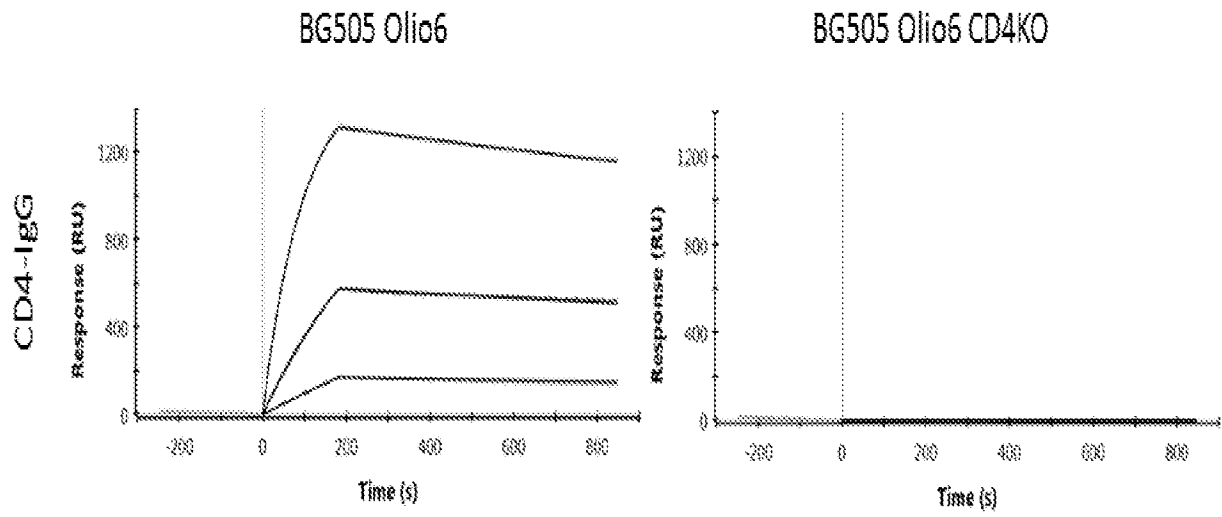


D. Binding to bnAbs assessed by ELISA, data as AUC

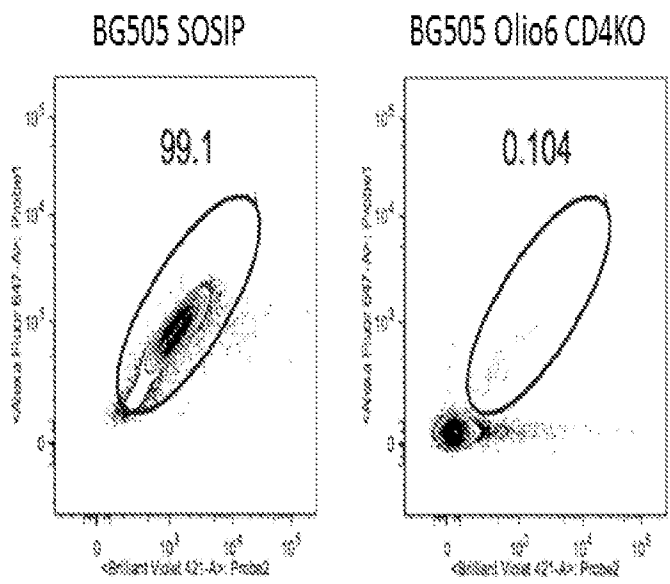


FIGS. 15C-15D

E. Binding to CD4 assessed by SPR



F. Binding to human CD4+ T cells assessed by cytometry



FIGS. 15E-15F

G. SECMAIS of BG505 MD39, BG505 Olio6, BG505 MD39C, BG505 MD39 CP1.2 GRSF4

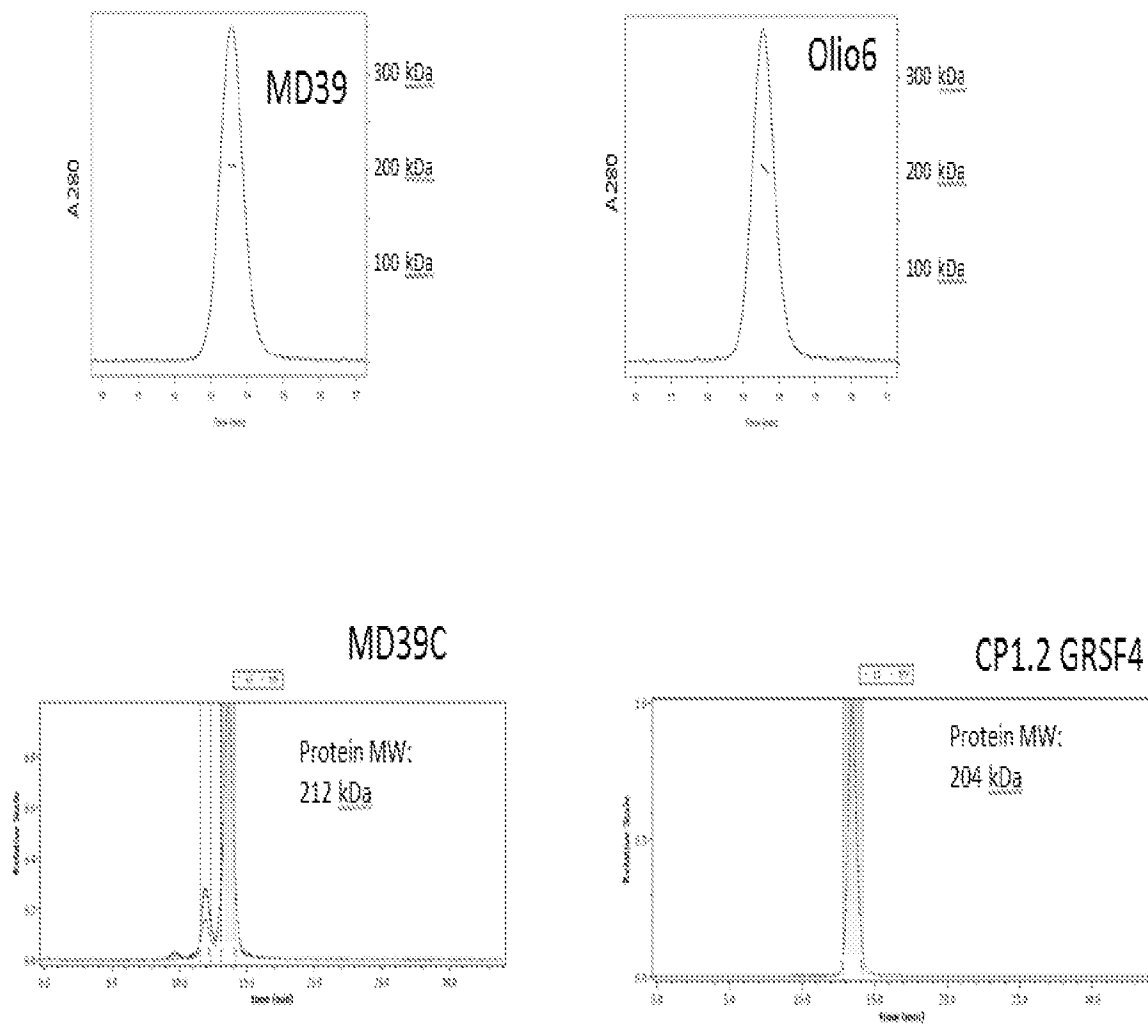


FIG. 15G

A. Binding KDs for bnAbs from SPR, and melting temperatures (TM)

Strain	Stabilization	Clade	V1V2 apex		N332 supersite		CD4 binding site			gp120-gp41 interface	TM
			PGDM 1400	PGT 145	PGT 121	PGT 128	VRC01	12A12	3BNC60	PGT 151	
BG505	SOSIP.664	A	6	15	22	17	131	97	137	10	66
	MD37	A				19		106	189		80
	MD39	A	12	20	21	18	124	67	121	11	77
	Olio6	A	18	31	15	11	90	40	119		76
	MD39C	A				14		71	144		82.5
	MD53	A	11	22							76
191084	MD39	A1	1.4	11	26	25	162	77	178		
AC10	MD39	B		9	169	33		277			60
	Olio6	B		53							59
001428	MD39	C		43	16	42	80	55	74		
B41	MD39	B		6							
ZM197	MD39	C		191							

FIG. 16A

B. DSC melting data for BG505 MD37, MD39, Olio6 and MD39C

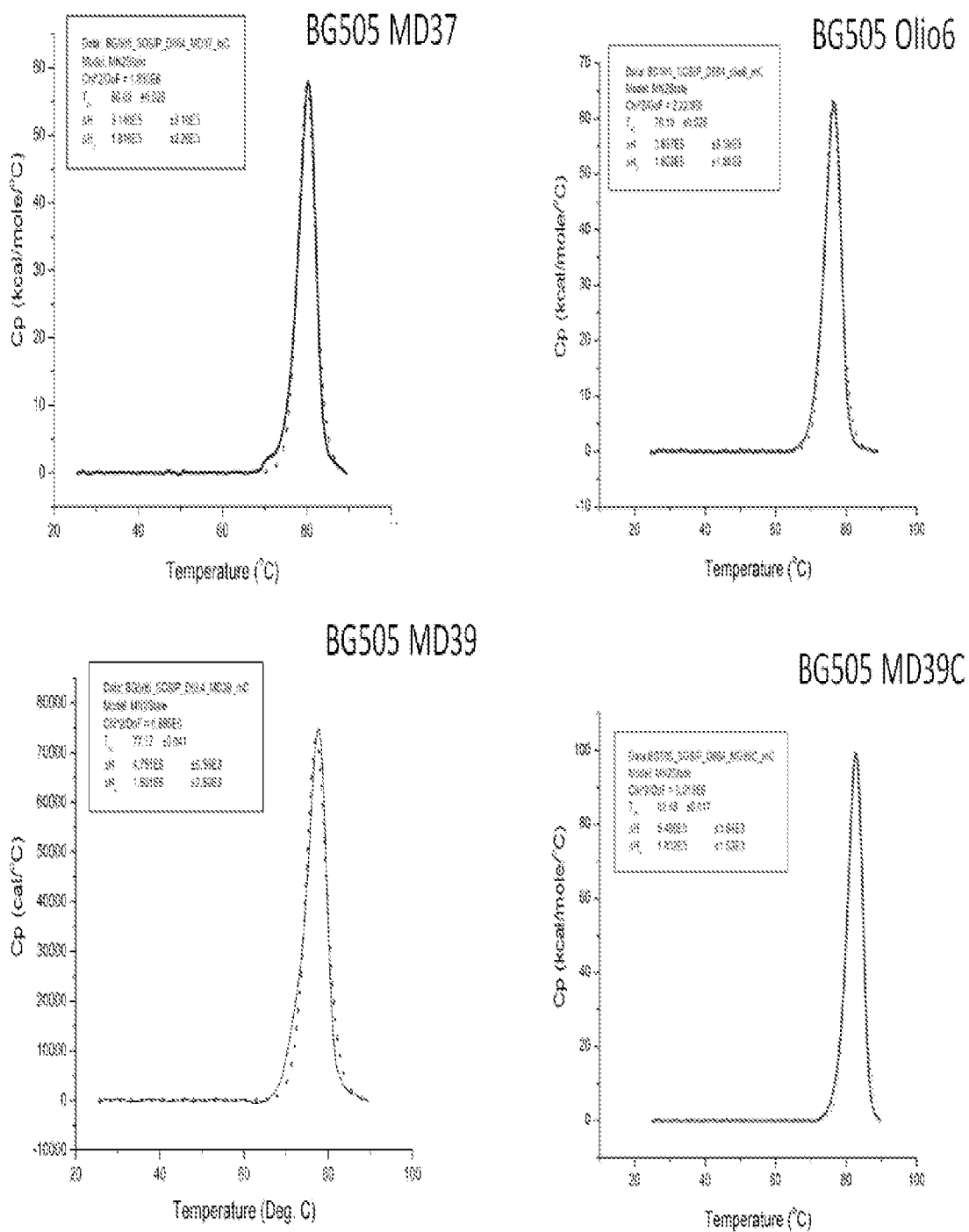


FIG. 16B

C. DSC melting data for BG505 MD53, MD39 CP1.1, MD39 CP1.2 and MD39 link14

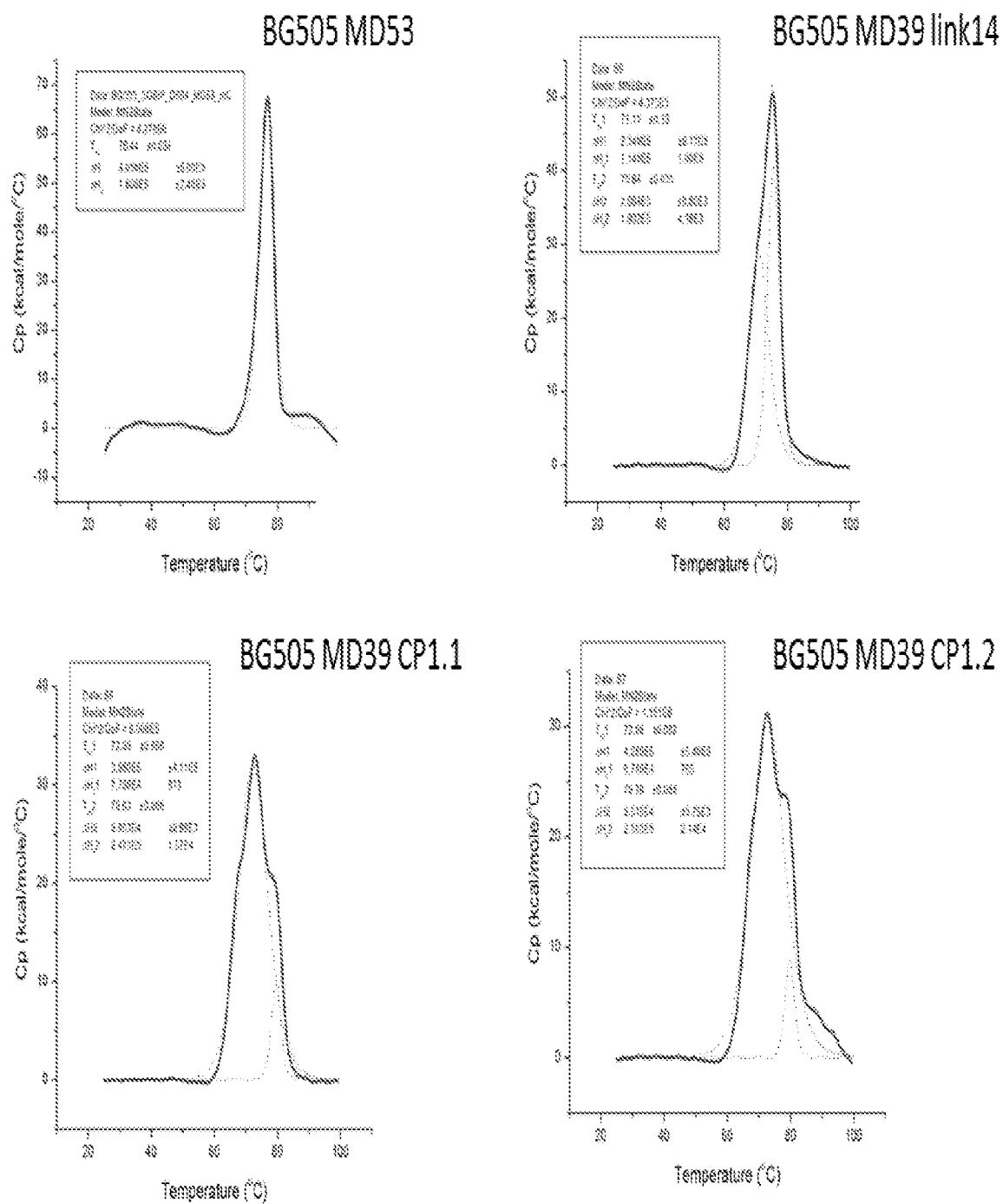
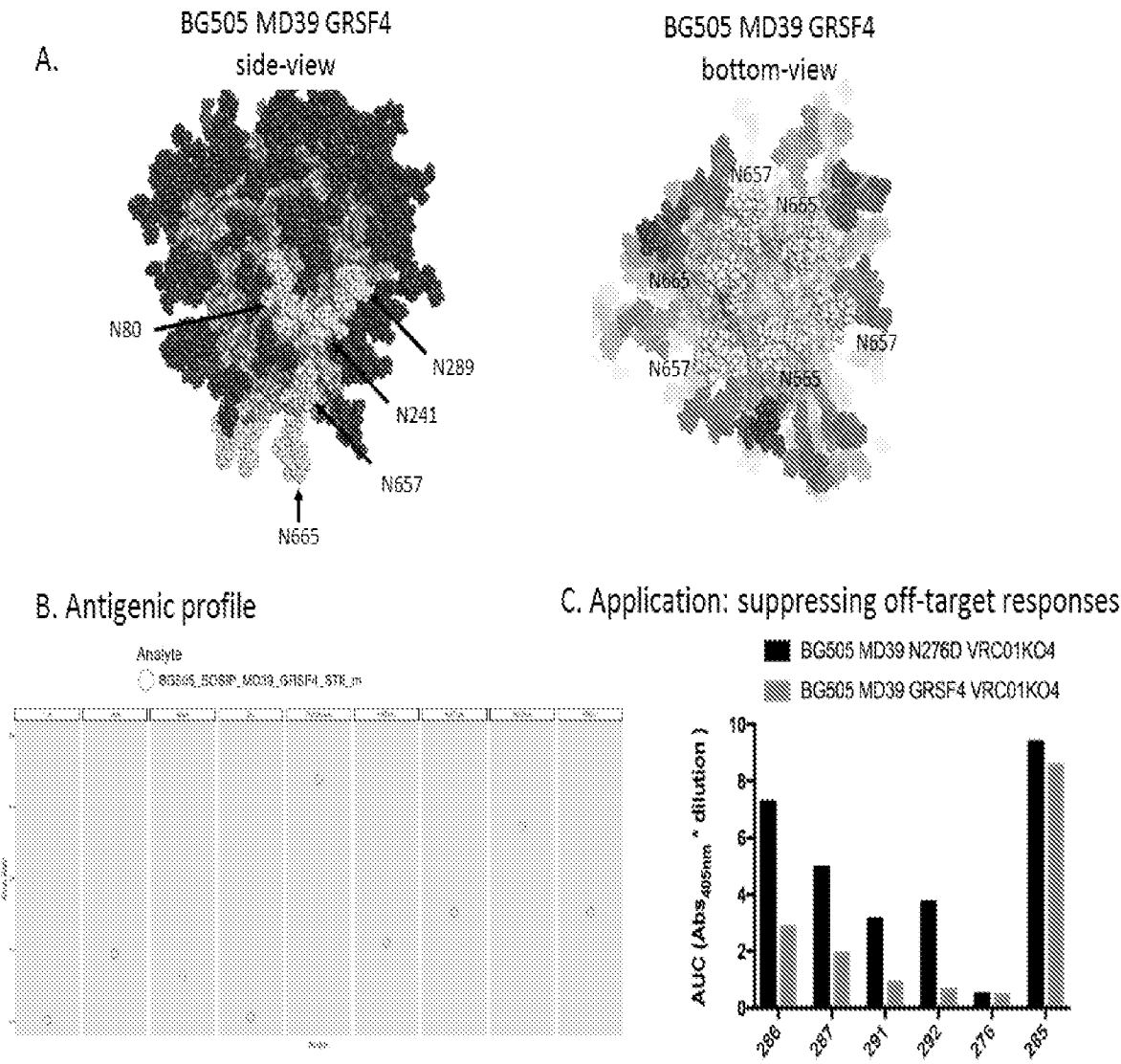
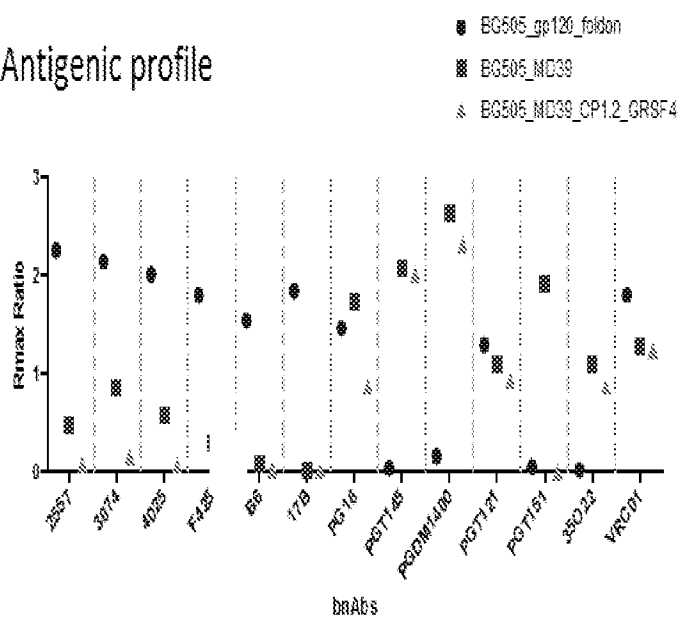


FIG. 16C

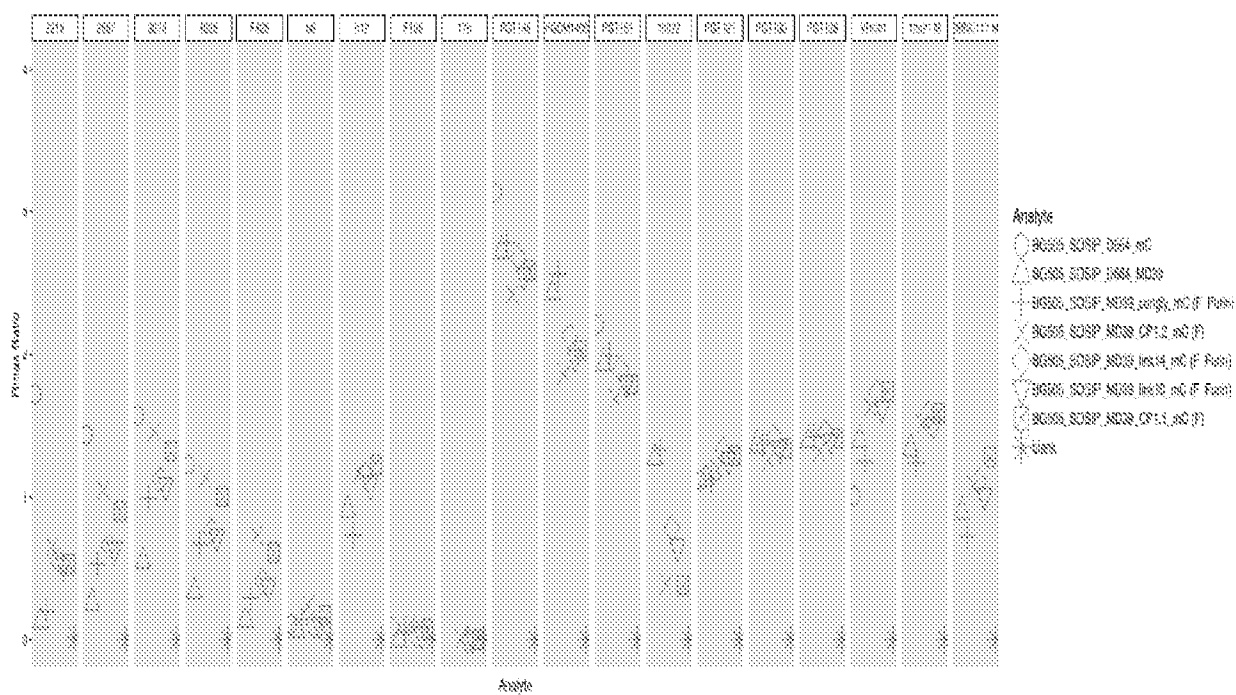


FIGS. 17A-17C

A. Antigenic profile

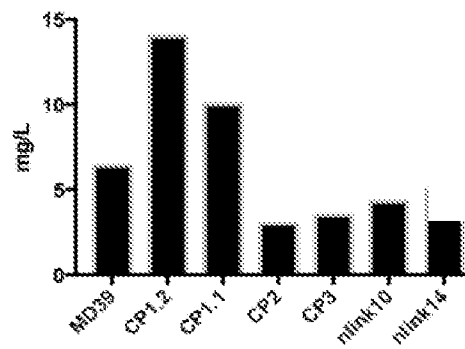


B. Antigenic profile



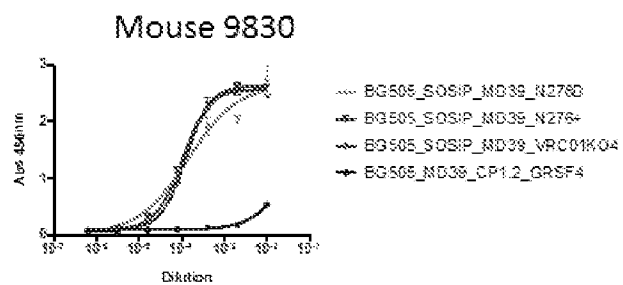
FIGS. 18A-18B

C. Expression levels



D. Example of Application: blocking off-target antibodies

- VRC01-gH KI mouse
- GT8, core-boost, 2x trimer boost



FIGS. 18C-18D

A. SECMAIS

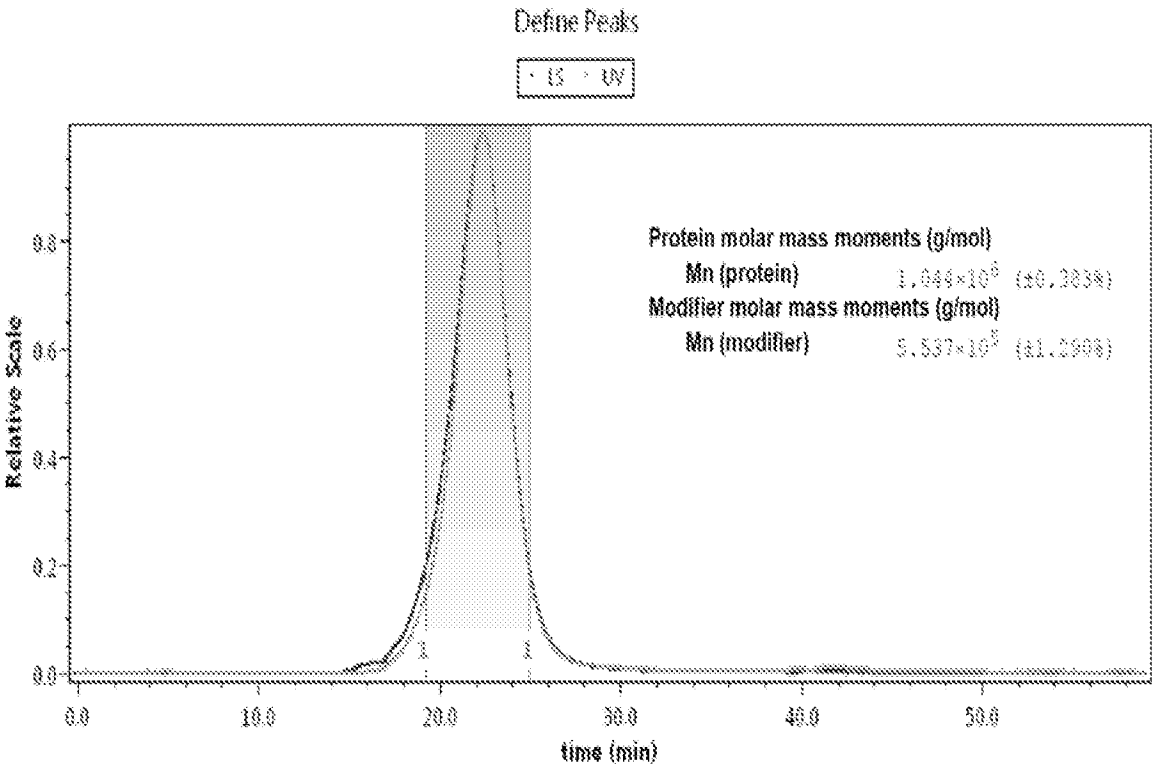


FIG. 19A

B. Negative stain EM

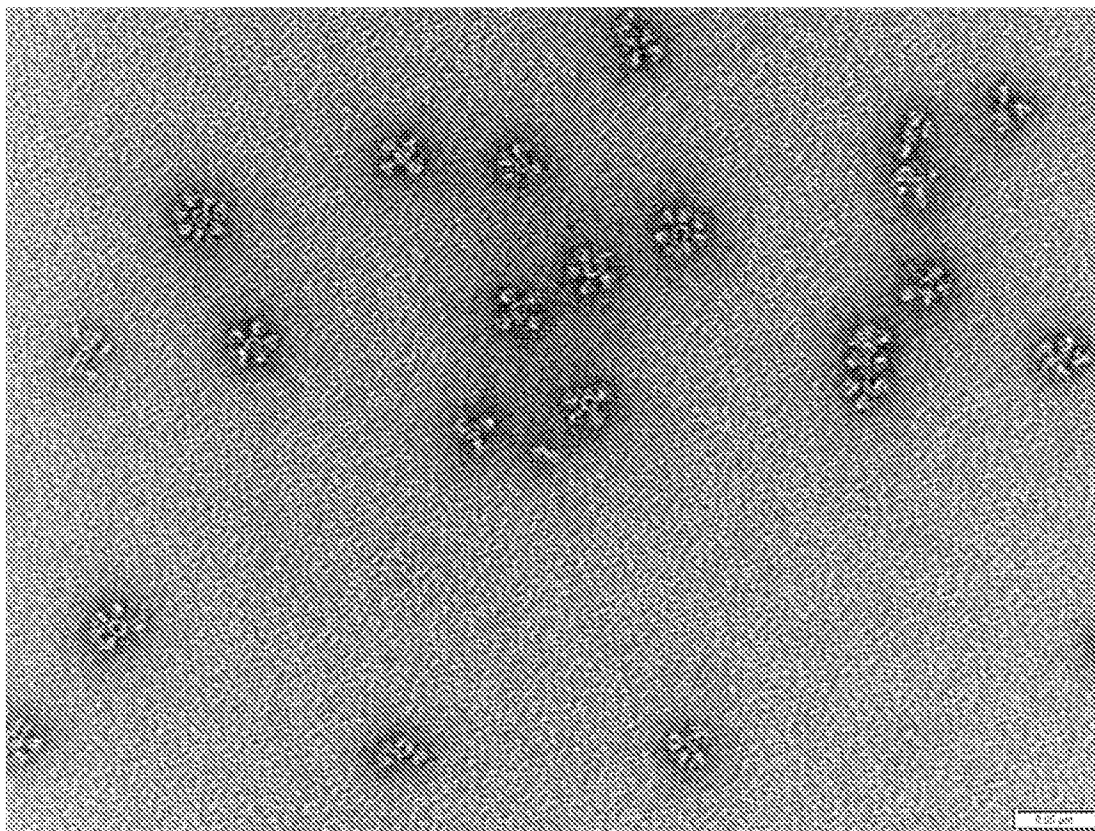


FIG. 19B

C. Negative stain EM

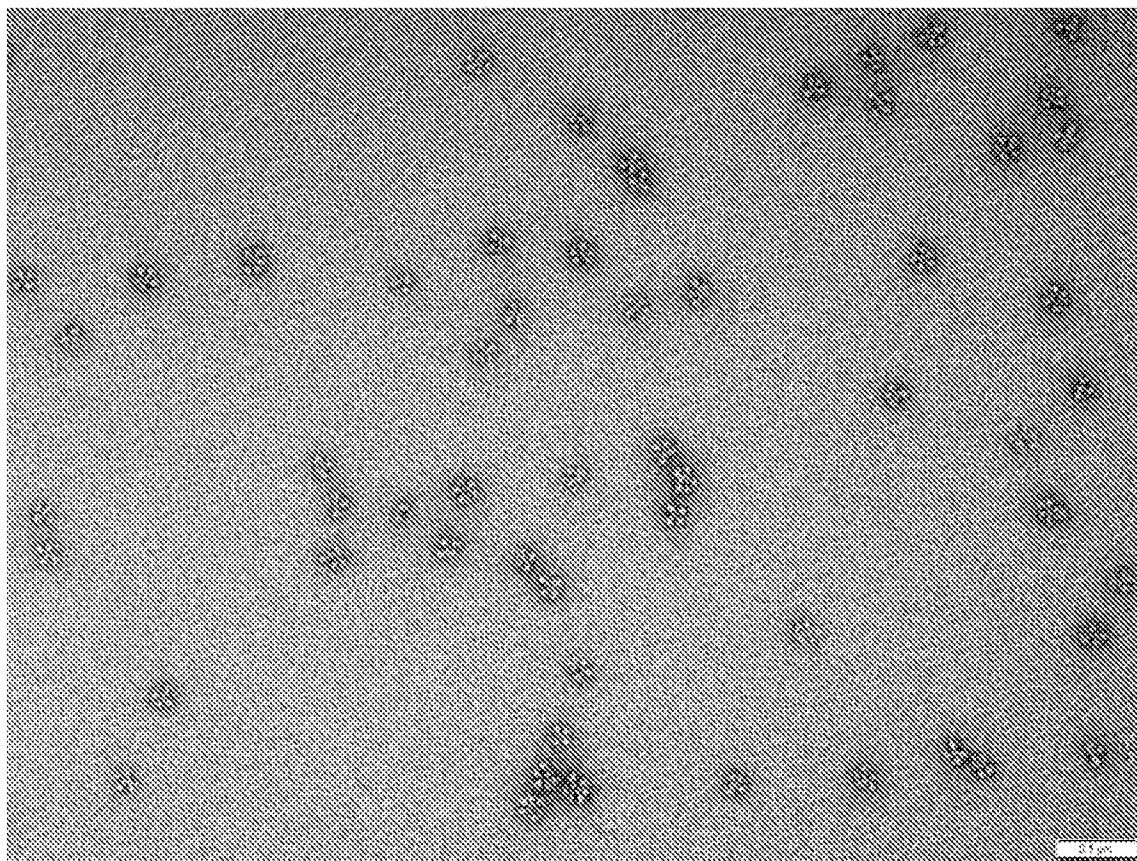


FIG. 19C

D. Antigenic profile (SPR)

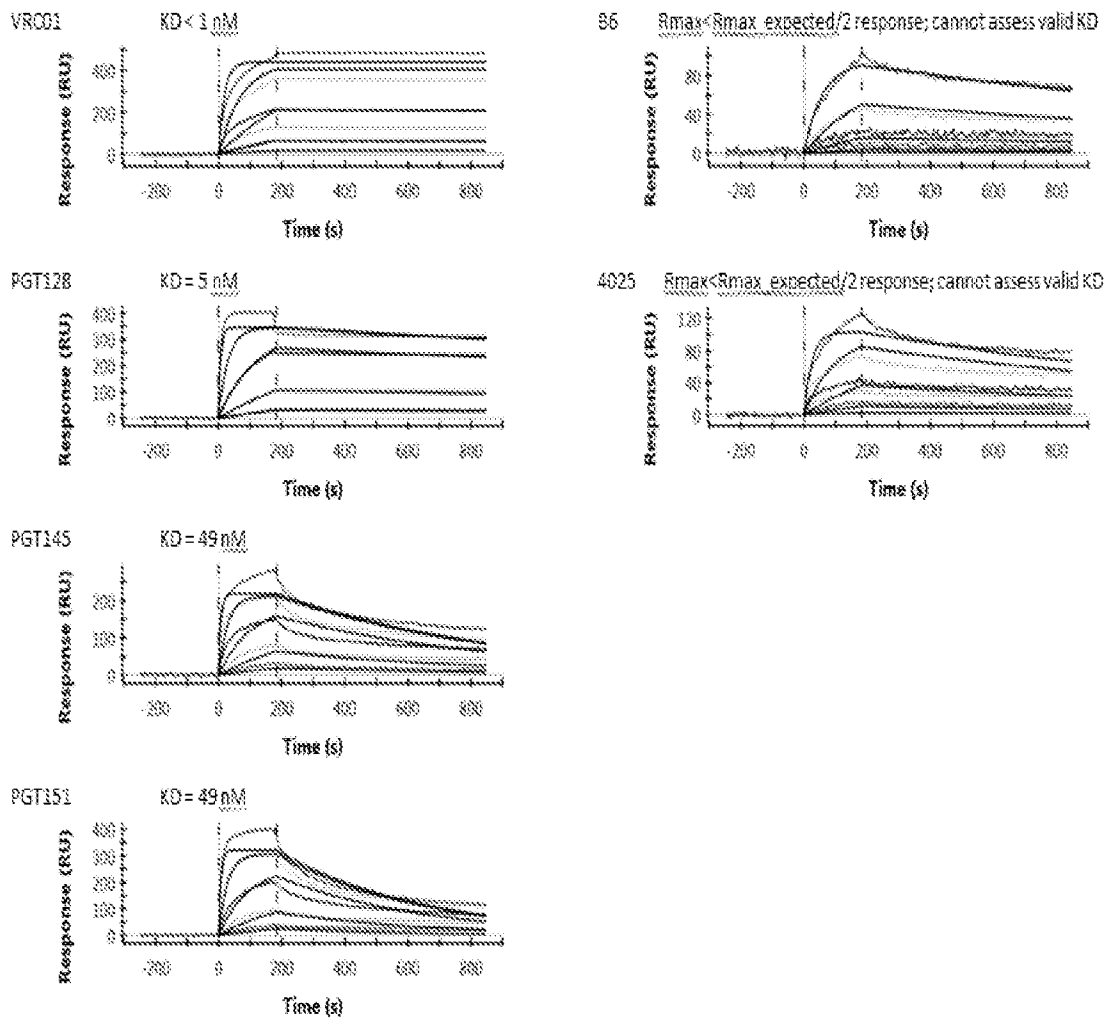


FIG. 19D

A. SECMAIS

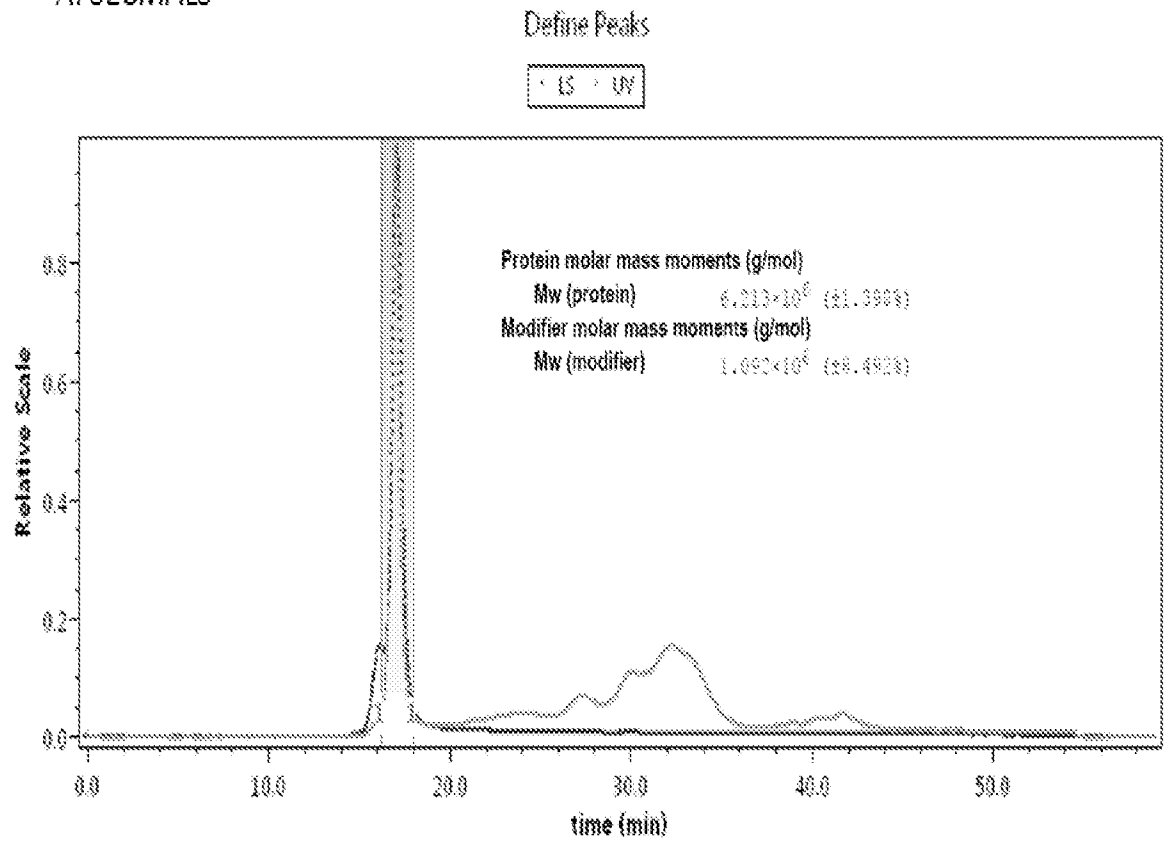


FIG. 20A

B. Negative stain EM

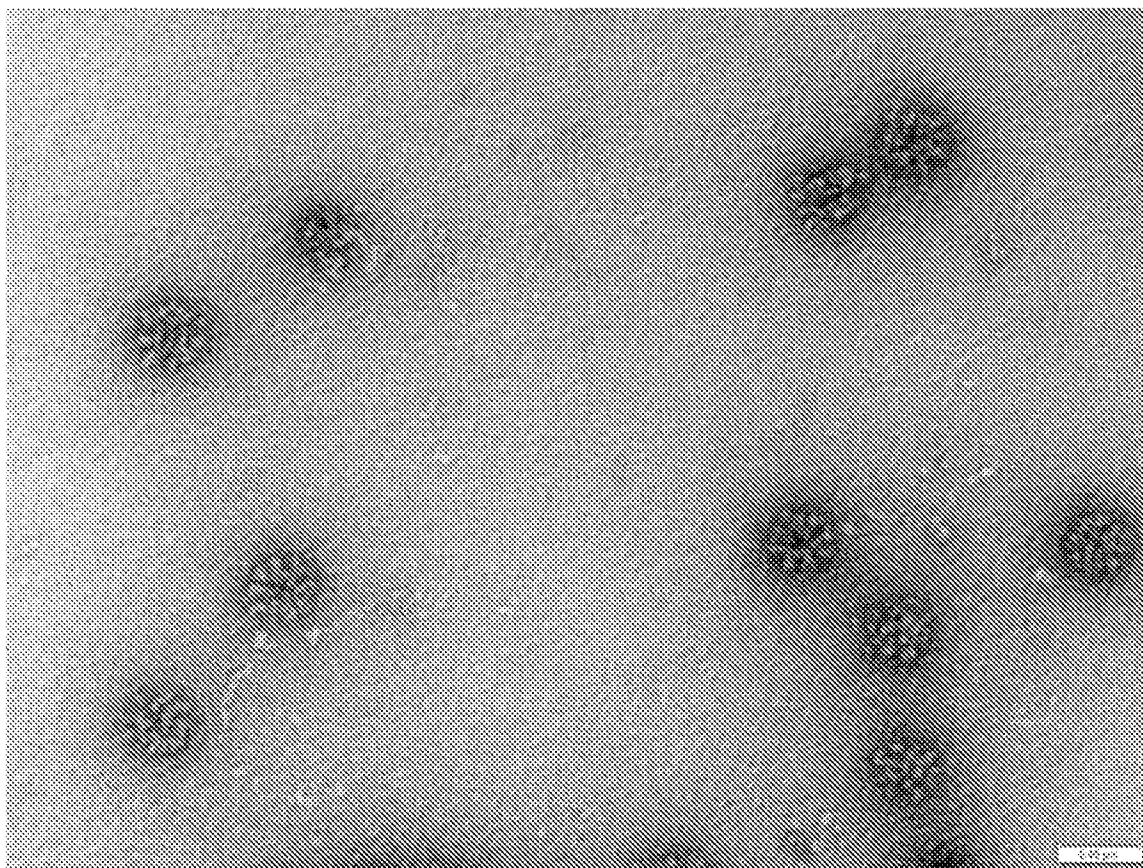


FIG. 20B

C. Negative stain EM

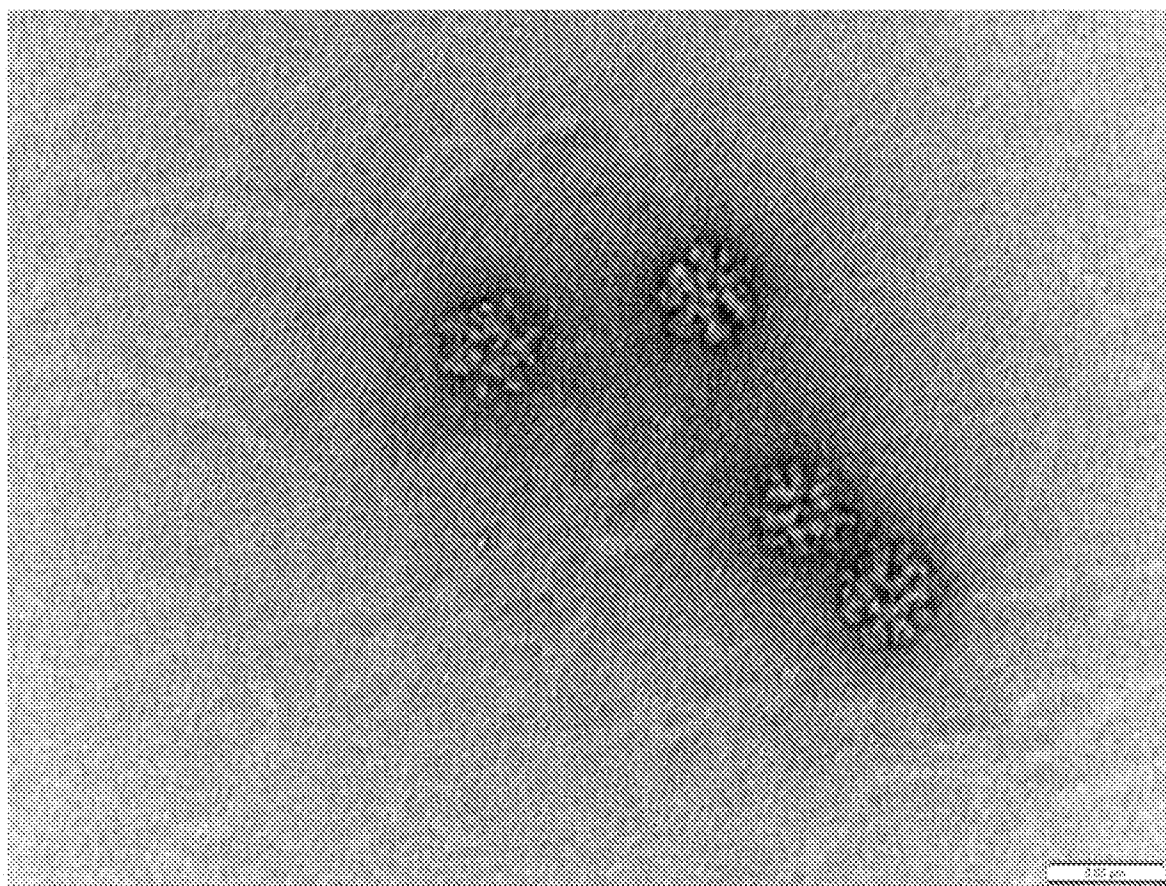


FIG. 20C

D. Antigenic profile (SPR)

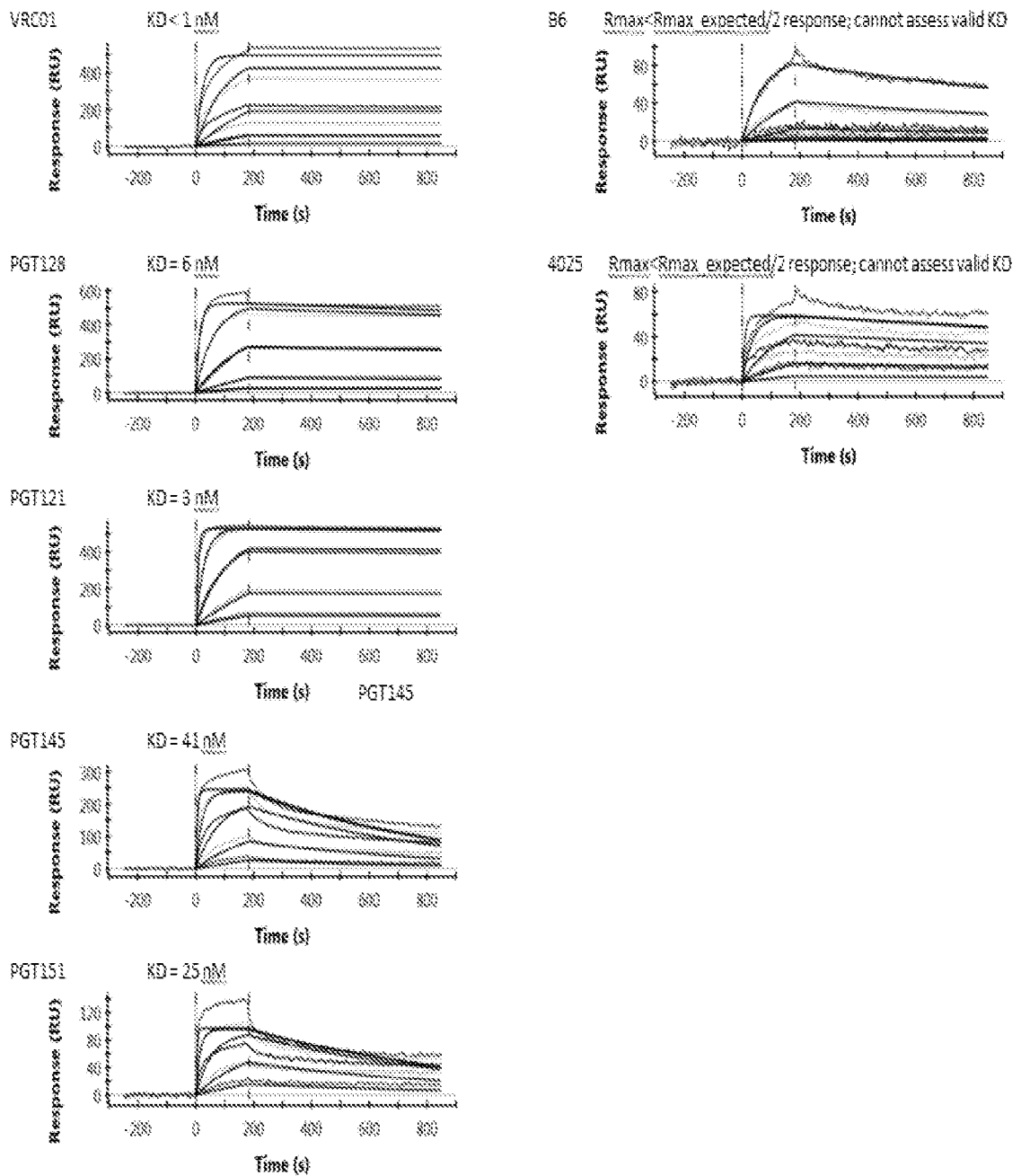
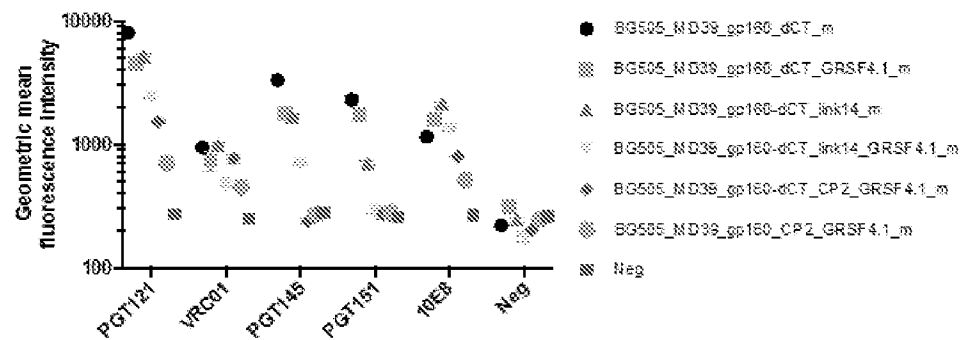
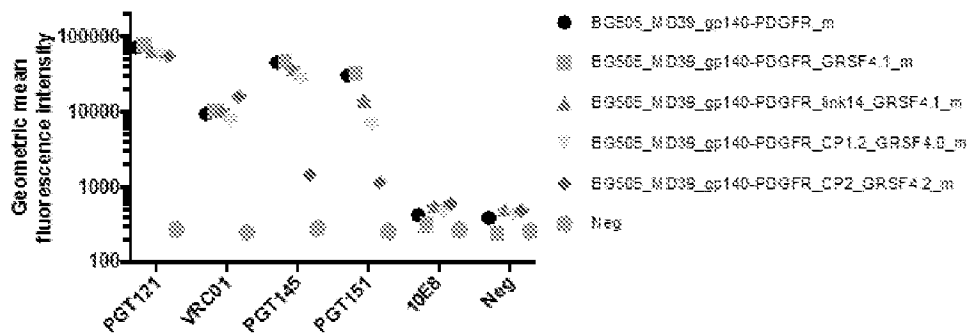


FIG. 20D

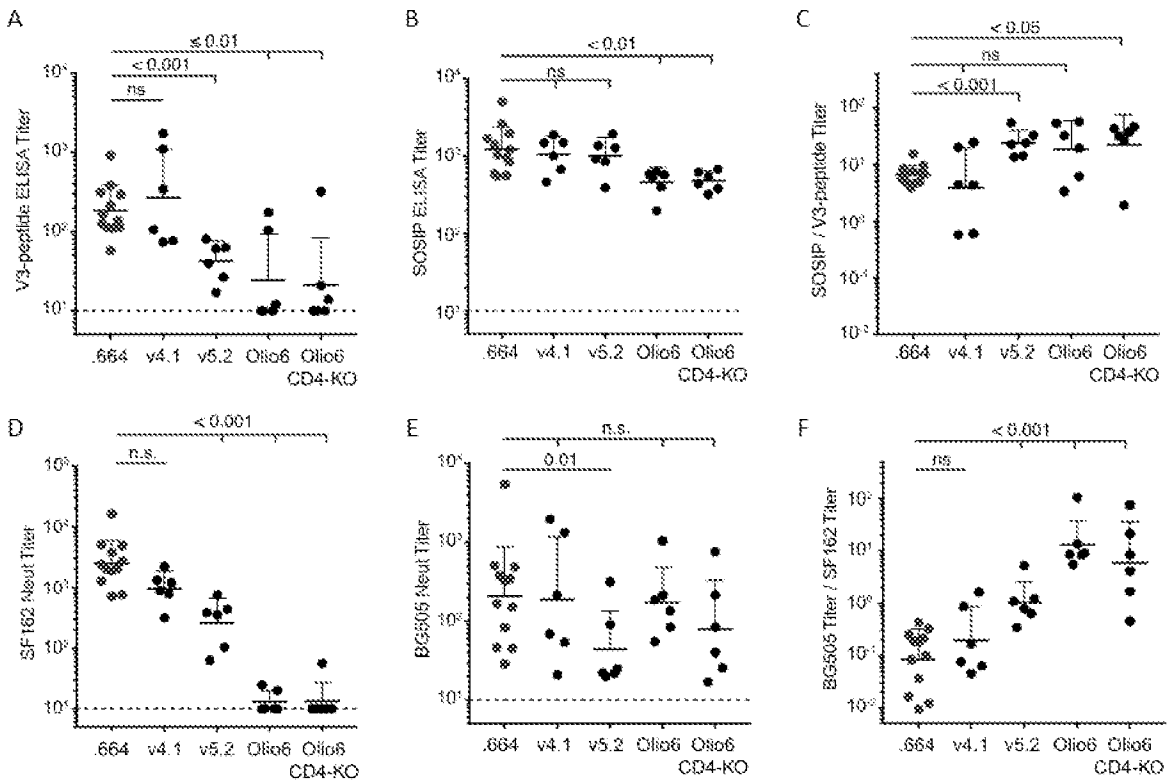
A. Binding of bNAbs to constructs with native transmembrane domains assessed by flow cytometry



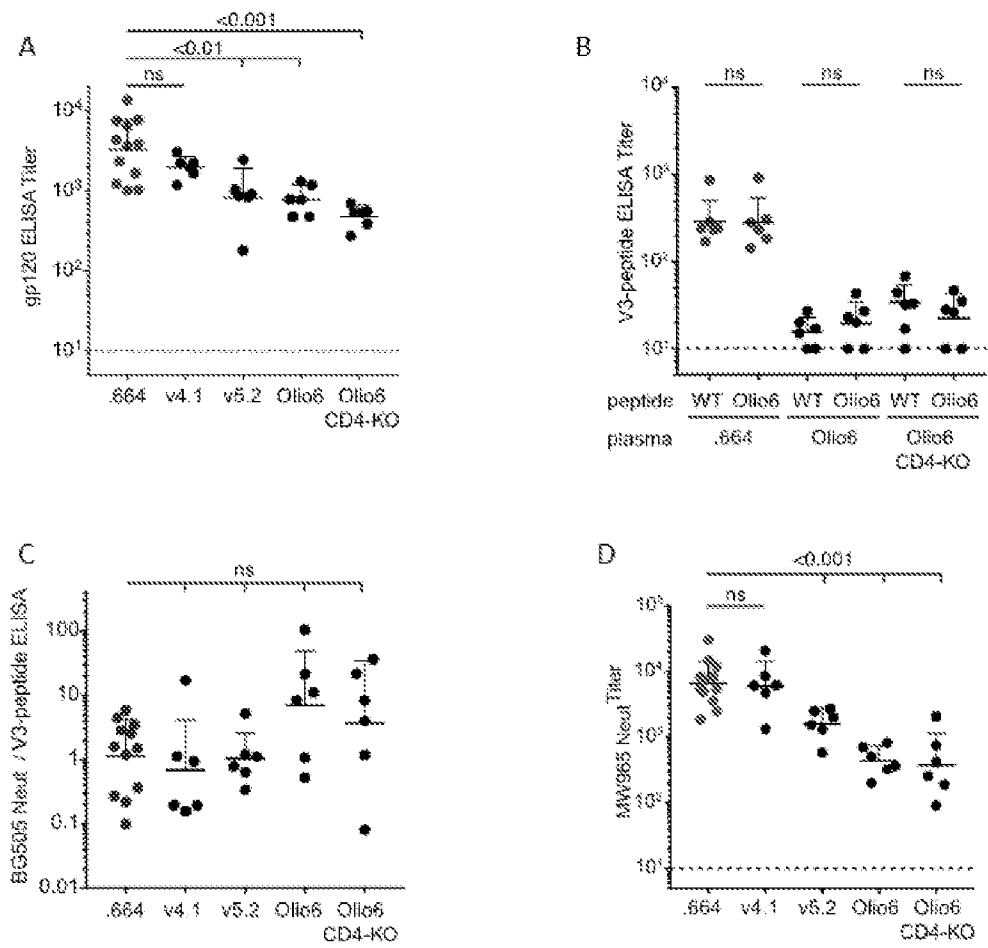
B. Binding of bNAbs to constructs anchored via a flexible linker and PDGFR transmembrane domain assessed by flow cytometry



FIGS. 21A-21B



FIGS. 22A-22F



FIGS. 23A-23D

42/42

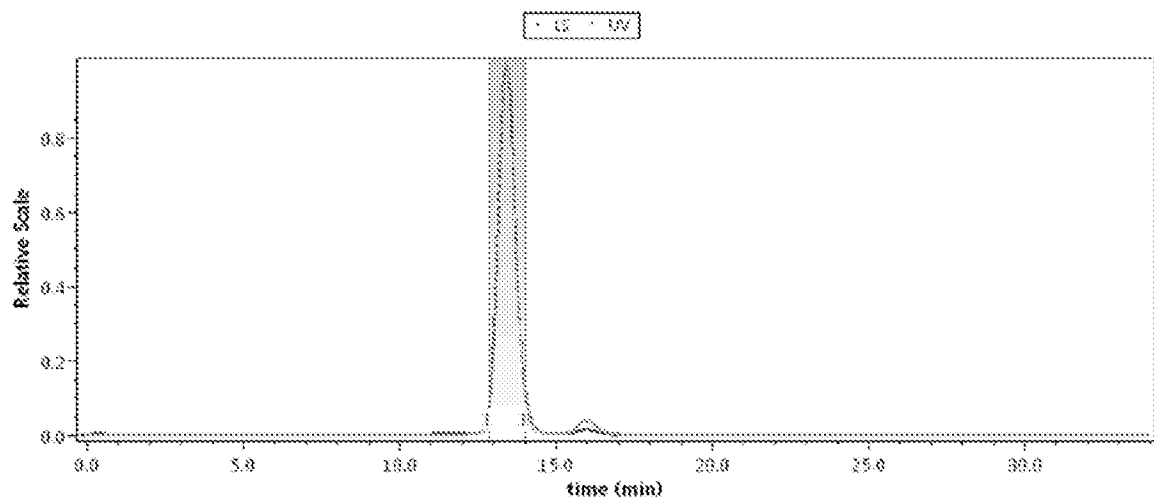
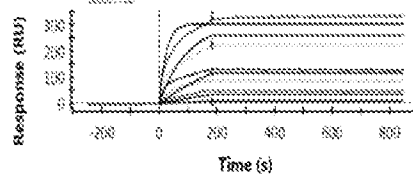
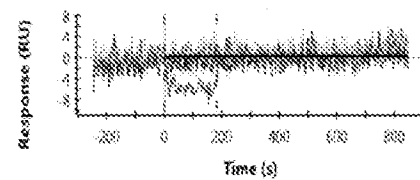
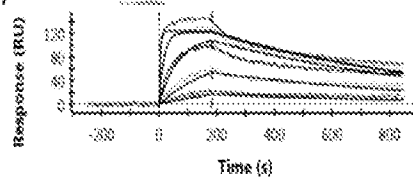


FIG. 24A

VRC01-Fab, $K_D < 1$ nM

4025-Fab, no detectable binding at 20uM max concentration.

PGT145-Fab, $K_D = 26$ nM

B6-Fab, no significant binding at 15uM max concentration

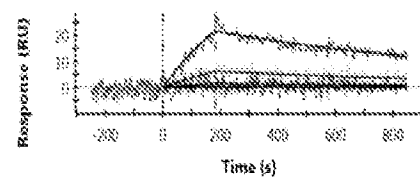
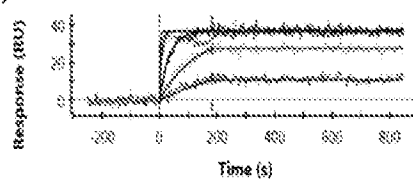
PGT121-Fab, $K_D < 1$ nM

FIG. 24B

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2017/023854

A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - A61K 39/21; A61K 39/40; A61K 39/42; A61K 45/06; A61P 31/18; C07K 16/10 (2017.01)

CPC - A61K 39/12; A61K 39/21; A61K 39/42; A61K 45/06; A61K 2039/64; C07K 16/1063; C12N 7/00; C12N 2740/16034 (2017.02)

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

See Search History document

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

USPC - 424/188.1; 530/350; 536/23.5 (keyword delimited)

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

See Search History document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2016/037154 A1 (THE UNITED STATES OF AMERICA, AS REPRESENTED BY THE SECRETARY, DEPARTMENT OF HEALTH & HUMAN SERVICES) 10 March 2016 (10.03.2016) entire document	1, 2, 15
A	WO 2015/134982 A1 (CORNELL UNIVERSITY) 11 September 2015 (11.09.2015) entire document	1-15
A	US 2014/0212458 A1 (INTERNATIONAL AIDS VACCINE INITIATIVE et al) 31 July 2014 (31.07.2014) entire document	1-15
A	US 2015/0238594 A1 (INTERNATIONAL AIDS VACCINE INITIATIVE et al) 27 August 2015 (27.08.2015) entire document	1-15
A	SANDERS et al. "A Next-Generation Cleaved, Soluble HIV-1 Env Trimer, BG505 SOSIP.664 gp140, Expresses Multiple Epitopes for Broadly Neutralizing but not non-Neutralizing Antibodies," PLoS Pathogens, 19 September 2013 (19.09.2013), Vol. 9, e1003618, Pgs. 1-20. entire document	1-15

☐ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

* Special categories of cited documents:

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search

27 June 2017

Date of mailing of the international search report

07 AUG 2017

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents
P.O. Box 1450, Alexandria, VA 22313-1450
Facsimile No. 571-273-8300

Authorized officer

Blaine R. Copenheaver

PCT Helpdesk: 571-272-4300
PCT OSP: 571-272-7774

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2017/023854

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

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

- a. ☐ forming part of the international application as filed:
 - ☐ in the form of an Annex C/ST.25 text file.
 - ☐ on paper or in the form of an image file.
- b. ☐ furnished together with the international application under PCT Rule 13*ter*. 1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.
- c. ☒ furnished subsequent to the international filing date for the purposes of international search only:
 - ☒ in the form of an Annex C/ST.25 text file (Rule 13*ter*. 1(a)).
 - ☐ on paper or in the form of an image file (Rule 13*ter*. 1(b) and Administrative Instructions, Section 713).

2. ☒ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, and 80-113 were searched.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2017/023854

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☒ Claims Nos.: 16-30
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.