



(51) International Patent Classification:
C07K 14/16 (2006.01)

(21) International Application Number:
PCT/US2020/057205

(22) International Filing Date:
23 October 2020 (23.10.2020)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/925,173 23 October 2019 (23.10.2019) US

(71) Applicants: **DUKE UNIVERSITY** [US/US]; 2812 Erwin Road, Suite 306, Durham, NC 27705 (US). **TRIAD NATIONAL SECURITY, LLC** [US/US]; LC/TAS, P.O. Box 1663, Los Alamos, NM 87545 (US).

(72) Inventors; and

(71) Applicants: **HAHN, Beatrice, H.** [US/US]; c/o The Trustees of The University of Pennsylvania, 3160 Chestnut Street, Suite 200, Philadelphia, PA 19104 (US). **RUSSELL, Ronnie, M.** [US/US]; c/o The Trustees of The University of Pennsylvania, 3160 Chestnut Street, Suite 200, Philadelphia, PA 19104 (US).

(72) Inventors: **SAUNDERS, Kevin, O.**; c/o Duke University, 2812 Erwin Road, Suite 306, Durham, NC 27705 (US). **HAYNES, Barton, F.**; 2812 Erwin Road, Suite 306, Durham, NC 27705 (US). **KORBER, Bette, T.**; c/o Triad National Security, LLC, Los Alamos National Laboratory, LC/TAS, P.O. Box 1663, Los Alamos, NM 87545 (US). **WAGH, Kshitij, G.**; c/o Triad National Security, LLC, Los Alamos National Laboratory, LC/TAS, P.O. Box 1663, Los Alamos, NM 87545 (US).

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

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

(54) Title: COMPOSITIONS COMPRISING V2 OPT HIV ENVELOPES

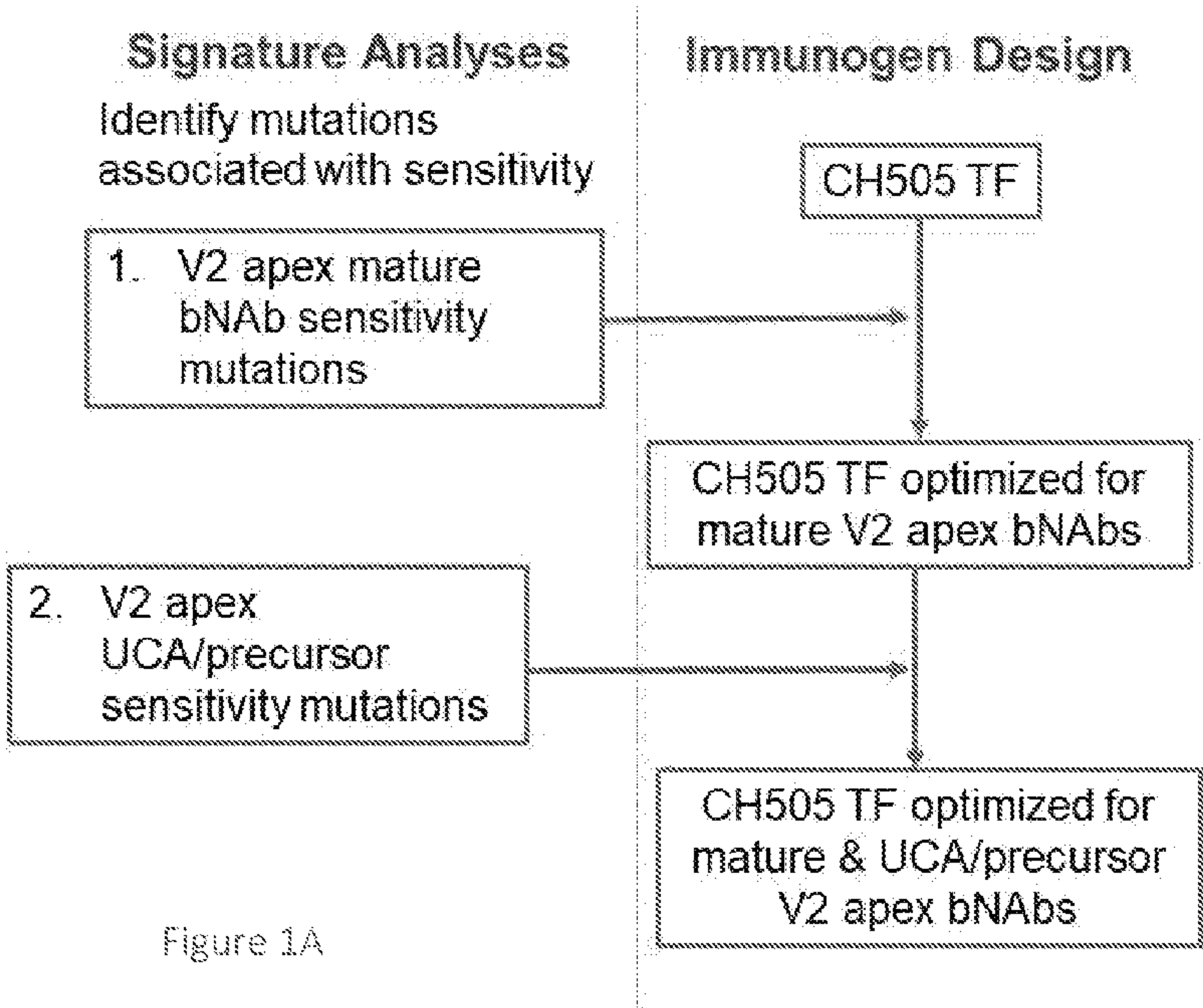


Figure 1A

(57) Abstract: In certain aspects the invention provides HIV-1 immunogens, including HIV-1 envelopes with optimized V2 loop for antibody induction.

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

Published:

- *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

Compositions Comprising V2 OPT HIV Envelopes

[0001] This application claims the benefit and priority of US Application Serial No. 62/925173 filed October 23, 2019 which content is incorporated by reference in its entirety.

[0002] This invention was made with government support under Center for HIV/AIDS Vaccine Immunology-Immunogen Design grant UM1-AI100645 and UM1-AI144371 from the NIH, NIAID, Division of AIDS. The government has certain rights in the invention.

[0003] The United States government has rights in this invention pursuant to Contract No. 89233218CNA000001 between the United States Department of Energy and Triad National Security, LLC for the operation of Los Alamos National Laboratory.

TECHNICAL FIELD

[0004] The present invention relates in general, to a composition suitable for use in inducing anti-HIV-1 antibodies, and, in particular, to immunogenic compositions comprising envelope proteins and nucleic acids to induce cross-reactive neutralizing antibodies and increase their breadth of coverage. The invention also relates to methods of inducing such broadly neutralizing anti-HIV-1 antibodies using such compositions.

BACKGROUND

[0005] The development of a safe and effective HIV-1 vaccine is one of the highest priorities of the scientific community working on the HIV-1 epidemic. While anti-retroviral treatment (ART) has dramatically prolonged the lives of HIV-1 infected patients, ART is not routinely available in developing countries.

SUMMARY OF THE INVENTION

[0006] In certain embodiments, the invention provides compositions and method for induction of immune response, for example cross-reactive (broadly) neutralizing Ab (bNAbs) induction.

[0007] In certain aspects, the invention provides CH505 envelope immunogens comprising optimized V2 loop, for example but not limited to initiate V1V2, and/or CD4 binding site

and/or Fusion Peptide unmutated common ancestor (UCA) broadly neutralizing antibody (bnAbs) precursors. In certain aspects the invention provides CH505 T/F envelope comprising optimized V2 loop.

[0008] In certain aspects the invention provides a recombinant HIV-1 envelope polypeptide optimized for V2 apex bnAb precursor targeting. In certain aspects the invention provides a recombinant HIV-1 envelope polypeptide optimized for V2 apex bnAb UCA/RUA (reverted unmutated ancestor) targeting. In non-limiting embodiments, the envelope is based on CH505 TF sequence. In non-limiting embodiments the envelope is any one of the envelopes from Table 1, Figures 3A-3C or Figures 4A-4B, wherein the polypeptide is a non-naturally occurring protomer designed to form an envelope trimer. In certain embodiments, the envelope is based on CH505 T/F envelope and comprises V2 binding antibody optimized sequence. In certain embodiments the envelope is designed to multimerize. In some embodiments the envelope sequence comprises a self-assembling protein, for example but not limited to ferritin. In other embodiments, the self-assembling protein is added via a sortase A reaction.

[0009] In certain embodiments, the optimized design has 11 mutations at positions 130, 161, 169, 170, 172, 316, 332, 335, 632, 644, and 717 as shown in Figure 1. In certain embodiments, the optimized design has additional five mutations at positions 161, 200, 305, 322, and 732 as shown in Figure 2.

[0010] In certain aspects, the invention provides a nucleic acid encoding the recombinant polypeptide of the invention. In non-limiting embodiments, the nucleic acid is mRNA. In non-limiting embodiments, the mRNA is modified. In non-limiting embodiments, the mRNA is administered as an LNP.

[0011] In certain aspects the invention provides a recombinant trimer comprising three identical protomers of an envelope of the invention, e.g. any envelope from Table 1, Figures 3A-3C or Figures 4A-4B.

[0012] In certain aspects, the invention provides an immunogenic composition comprising the recombinant trimer and a carrier, wherein the trimer comprises three identical protomers of an HIV-1 envelope listed in Table 1, Figures 3A-3C or Figures 4A-4B.

[0013] In certain aspects, the invention provides an immunogenic composition comprising nucleic acid encoding the recombinant HIV-1 envelope and a carrier.

[0014] In certain aspects, the immunogenic compositions of the invention are further comprise an adjuvant.

[0015] In certain embodiments, the nucleic acid of the invention is operably linked to a promoter, and wherein in certain embodiment the nucleic acid is inserted in an expression vector.

[0016] In certain aspects, the invention provides a method of inducing an immune response in a subject comprising administering a composition comprising any suitable form of a nucleic acid(s) of the invention or the recombinant polypeptide of the invention in an amount sufficient to induce an immune response.

[0017] In certain embodiments, the nucleic acid encodes a gp120 envelope, gp120D8 envelope, a gp140 envelope (gp140C, gp140CF, gp140CFI) as a soluble or stabilized protomer of a SOSIP trimer, a gp145 envelope, a gp150 envelope, a transmembrane bound envelope, or an envelope designed to multimerize.

[0018] In certain embodiments, the recombinant polypeptide is a gp120 envelope, gp120D8 envelope, a gp140 envelope (gp140C, gp140CF, gp140CFI) as a soluble or stabilized protomer of a SOSIP trimer, a gp145 envelope, a gp150 envelope, a transmembrane bound envelope, or an envelope designed to multimerize. In non-limiting embodiments, the envelope multimerizes via ferritin.

[0019] In certain embodiments, the composition further comprises an adjuvant.

[0020] In certain embodiments, the methods further comprise administering an agent which modulates host immune tolerance.

[0021] In certain embodiments, the recombinant polypeptide administered is multimerized in a liposome or nanoparticle.

[0022] In certain embodiments, the methods further comprise administering one or more additional HIV-1 immunogens to induce a T cell response.

[0023] In certain aspects, the invention provides a composition comprising a nanoparticle and a carrier, wherein the nanoparticle comprises any one of the envelopes of the invention.

[0024] In certain embodiments, the nanoparticle is ferritin self-assembling nanoparticle.

[0025] In certain aspects, the invention provides a composition comprising a nanoparticle and a carrier, wherein the nanoparticle comprises any one of the trimers of the invention.

[0026] In certain embodiments, the nanoparticle is a ferritin self-assembling nanoparticle.

[0027] In certain embodiments, the nanoparticle comprises multimers of trimers. In certain embodiments, the nanoparticle comprises 1-8 trimers.

[0028] In certain aspects, the invention provides methods of inducing an immune response in a subject comprising administering a sufficient amount to induce an immune response of an

immunogenic composition comprising any one of the recombinant envelopes described herein or compositions described herein. In certain embodiments, the composition is administered as a prime. In certain embodiments, the composition is administered as a boost.

[0029] In certain embodiments, the invention provides a nucleic acid encoding any of the recombinant envelopes of the invention.

[0030] In certain embodiments, the invention provides a composition comprising the nucleic acid of the invention and a carrier.

[0031] In certain embodiments, the invention provides a method of inducing an immune response in a subject comprising administering an immunogenic composition comprising the nucleic acid or the composition comprising the nucleic acids of the invention.

[0032] In certain embodiments, the compositions contemplate nucleic acid, as DNA and/or RNA, or proteins immunogens either alone or in any combination. In certain embodiments, the methods contemplate genetic, as DNA and/or RNA, immunization either alone or in combination with envelope protein(s).

[0033] In certain embodiments the nucleic acid encoding an envelope is operably linked to a promoter inserted an expression vector. In certain aspects the compositions comprise a suitable carrier. In certain aspects the compositions comprise a suitable adjuvant.

[0034] In certain embodiments the induced immune response includes induction of antibodies, including but not limited to autologous and/or cross-reactive (broadly) neutralizing antibodies against HIV-1 envelope. Various assays that analyze whether an immunogenic composition induces an immune response, and the type of antibodies induced are known in the art and are also described herein.

[0035] In certain aspects the invention provides a nucleic acid sequence encoding any of the polypeptides of the invention, wherein the nucleic acid is operably linked to a promoter. In certain aspects the invention provides a nucleic acid consisting essentially of a nucleic acid sequence encoding any of the polypeptides of the invention, wherein the nucleic acid is operably linked to a promoter. In certain aspects the invention provides an expression vector comprising any of the nucleic acid sequences of the invention, wherein the nucleic acid is operably linked to a promoter. In certain aspects the invention provides an expression vector comprising a nucleic acid sequence encoding any of the polypeptides of the invention, wherein the nucleic acid is operably linked to a promoter. In certain aspects the invention provides an expression vector consisting essentially a nucleic acid sequence encoding any of the polypeptides of the invention, wherein the nucleic acid is operably linked to a promoter.

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

[0036] In certain aspects the invention provides a composition comprising at least one of the nucleic acid sequences of the invention. In certain aspects the invention provides a composition comprising any one of the nucleic acid sequences of invention. In certain aspects the invention provides a composition comprising at least one nucleic acid sequence encoding any one of the polypeptides of the invention.

[0037] In certain aspects the invention provides a composition comprising at least one nucleic acid encoding an HIV-1 envelope of the invention.

[0038] In certain embodiments, the compositions and methods employ an HIV-1 envelope as polypeptide instead of a nucleic acid sequence encoding the HIV-1 envelope. In certain embodiments, the compositions and methods employ an HIV-1 envelope as polypeptide, a nucleic acid sequence encoding the HIV-1 envelope, or a combination thereof. In certain embodiments, the polypeptides are recombinantly produced.

[0039] The envelope used in the compositions and methods of the invention can be a gp160, gp150, gp145, gp140, gp120, gp41, or N-terminal deletion variants thereof as described herein, cleavage resistant variants thereof as described herein, or codon optimized sequences thereof. In certain embodiments the composition comprises envelopes as trimers. In certain embodiments, envelope proteins are multimerized, for example trimers are attached to a particle such that multiple copies of the trimer are attached and the multimerized envelope is prepared and formulated for immunization in a human. In certain embodiments, the compositions comprise envelopes, including but not limited to trimers as particulate, high-density array on liposomes or other particles, for example but not limited to nanoparticles. In some embodiments, the trimers are in a well ordered, near native like or closed conformation. In some embodiments the trimer compositions comprise a homogenous mix of native like

trimers. In some embodiments the trimer compositions comprise at least 65%, 70%, 75%, 80%, 85%, 90%, 95% native like trimers.

[0040] The polypeptide contemplated by the invention can be a polypeptide comprising any one of the polypeptides described herein. The polypeptide contemplated by the invention can be a polypeptide consisting essentially of any one of the polypeptides described herein. The polypeptide contemplated by the invention can be a polypeptide consisting of any one of the polypeptides described herein. In certain embodiments, the polypeptide is recombinantly produced. In certain embodiments, the polypeptides and nucleic acids of the invention are suitable for use as an immunogen, for example to be administered in a human subject.

[0041] In certain embodiments the envelope is any of the forms of HIV-1 envelope. In certain embodiments the envelope is a gp120, gp140, gp145 (i.e. with a transmembrane), gp150 envelope. In certain embodiments, gp140 is designed to form a stable trimer. In certain embodiments envelope protomers from a trimer which is not a SOSIP trimer. In certain embodiment the trimer is a SOSIP based trimer wherein each protomer comprises additional modifications. In certain embodiments, envelope trimers are recombinantly produced. In certain embodiments, envelope trimers are purified from cellular recombinant fractions by antibody binding and reconstituted in lipid comprising formulations. See for example WO2015/127108 titled "Trimeric HIV-1 envelopes and uses thereof" which content is herein incorporated by reference in its entirety. In certain embodiments the envelopes of the invention are engineered and comprise non-naturally occurring modifications.

[0042] In certain embodiments, the envelope is in a liposome. In certain embodiments the envelope comprises a transmembrane domain with a cytoplasmic tail embedded in a liposome. In certain embodiments, the nucleic acid comprises a nucleic acid sequence which encodes a gp120, gp140, gp145, gp150, gp160.

[0043] In certain embodiments, where the nucleic acids are operably linked to a promoter and inserted in a vector, the vectors are any suitable vector. Non-limiting examples include, VSV, replicating rAdenovirus type 4, MVA, Chimp adenovirus vectors, pox vectors, and the like. In certain embodiments, the nucleic acids are administered in NanoTaxi block polymer nanospheres. In certain embodiments, the composition and methods comprise an adjuvant. Non-limiting examples include, AS01 B, AS01 E, gla/SE, alum, Poly I poly C (poly IC), polyIC/long chain (LC) TLR agonists, TLR7/8 and 9 agonists, or a combination of TLR7/8 and TLR9 agonists (see Moody et al. (2014) J. Virol. March 2014 vol. 88 no. 6 3329-3339), or any other adjuvant. Non-limiting examples of TLR7/8 agonist include TLR7/8 ligands,

Gardiquimod, Imiquimod and R848 (resiquimod). A non-limiting embodiment of a combination of TLR7/8 and TLR9 agonist comprises R848 and oCpG in STS (see Moody et al. (2014) J. Virol. March 2014 vol. 88 no. 6 3329-3339).

[0044] In certain aspects the invention provides a cell comprising a nucleic acid encoding any one of the envelopes of the invention suitable for recombinant expression. In certain aspects, the invention provides a clonally derived population of cells encoding any one of the envelopes of the invention suitable for recombinant expression. In certain aspects, the invention provides a stable pool of cells encoding any one of the envelopes of the invention suitable for recombinant expression.

[0045] In certain aspects, the invention provides a recombinant HIV-1 envelope polypeptide listed in Table 1. In certain embodiments, the polypeptide is a non-naturally occurring protomer designed to form an envelope trimer. The invention also provides nucleic acids encoding these recombinant polypeptides. Non-limiting examples of amino acids and nucleic acid of such protomers are shown in Figures 3A-5E.

[0046] In certain aspects the invention provides a recombinant trimer comprising three identical protomers of an envelope from Table 1. In certain aspects the invention provides an immunogenic composition comprising the recombinant trimer and a carrier, wherein the trimer comprises three identical protomers of an HIV-1 envelope listed in Table 1. In certain aspects the invention provides an immunogenic composition comprising a nucleic acid encoding these recombinant HIV-1 envelope and a carrier.

[0047] In certain aspects the invention provides nucleic acids encoding HIV-1 envelopes for immunization wherein the nucleic acid encodes a gp120 envelope, gp120D8 envelope, a gp140 envelope (gp140C, gp140CF, gp140CFI) as soluble or stabilized protomer of a SOSIP trimer, a gp145 envelope, a gp150 envelope, or a transmembrane bound envelope.

[0048] In certain aspects the invention provides a selection of HIV-1 envelopes for immunization wherein the HIV-1 envelope is a gp120 envelope or a gp120D8 variant. In certain embodiments a composition for immunization comprises protomers that form stabilized SOSIP trimers.

[0049] In certain embodiments, the compositions for use in immunization further comprise an adjuvant.

[0050] In certain embodiments, wherein the compositions comprise a nucleic acid, the nucleic acid is operably linked to a promoter, and could be inserted in an expression vector.

[0051] In one aspect the invention provides a composition for a prime boost immunization regimen comprising one or more envelopes from Table 1, wherein the polypeptide is a non-naturally occurring protomer designed to form an envelope trimer, wherein the envelope is a prime or boost immunogen. In one aspect the invention provides a composition for a prime boost immunization regimen comprising one or more envelopes of the invention.

[0052] In certain aspects the invention provides methods of inducing an immune response in a subject comprising administering a composition comprising a polypeptide and/or any suitable form of a nucleic acid(s) encoding an HIV-1 envelope(s) in an amount sufficient to induce an immune response.

[0053] In certain embodiments, the nucleic acid encodes a gp120 envelope, gp120D8 envelope, a gp140 envelope (gp140C, gp140CF, gp140CFI) as soluble or stabilized protomer of a SOSIP trimer, a gp145 envelope, a gp150 envelope, or a transmembrane bound envelope. In certain embodiments, the polypeptide is gp120 envelope, gp120D8 envelope, a gp140 envelope (gp140C, gp140CF, gp140CFI) as soluble or stabilized protomer of a SOSIP trimer, a gp145 envelope, a gp150 envelope, or a transmembrane bound envelope.

[0054] In certain embodiments, the methods comprise administering an adjuvant. In certain embodiments, the methods comprise administering an agent which modulates host immune tolerance. In certain embodiments, the administered polypeptide is multimerized in a liposome or nanoparticle. In certain embodiments, the methods comprise administering one or more additional HIV-1 immunogens to induce a T cell response. Non-limiting examples include gag, nef, pol, etc.

[0055] In certain aspects, the invention provides a recombinant HIV-1 Env ectodomain trimer, comprising three gp120-gp41 protomers comprising a gp120 polypeptide and a gp41 ectodomain, wherein each protomer is the same and each protomer comprises portions from envelope BG505 HIV-1 strain and gp120 polypeptide portions from a CH505 HIV-1 strain and stabilizing mutations A316W and E64K. In certain embodiments, the trimer is stabilized in a prefusion mature closed conformation, and wherein the trimer does not comprise non-natural disulfide bond between cysteine substitutions at positions 201 and 433 of the HXB2 reference sequence. Non-limited examples of envelopes contemplated as trimers are listed in Table 1. In some embodiments, the amino acid sequence of one monomer comprised in the trimer is shown in Figure 3-5. In some embodiments, the trimer is immunogenic. In some embodiments the trimer binds to any one of the antibodies PGT145, PGT151, CH103UCA,

CH103, VRC01, PGT128, or any combination thereof. In some embodiments the trimer does not bind to antibody 19B and/or 17B.

[0056] In certain aspects, the invention provides a pharmaceutical composition comprising any one of the recombinant trimers of the invention. In certain embodiments the compositions comprising trimers are immunogenic. The percent trimer in such immunogenic compositions could vary. In some embodiments the composition comprises 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% stabilized trimer.

[0057] In certain embodiments, the envelope comprise ferritin. In certain embodiments, the inventive designs comprise modifications, including without limitation linkers between the envelope and ferritin designed to optimize ferritin nanoparticle assembly.

[0058] In certain aspects, the invention provides a composition comprising any one of the inventive envelopes or nucleic acid sequences encoding the same. In certain embodiments, the nucleic acid is mRNA. In certain embodiments, the mRNA is comprised in a lipid nanoparticle (LNP).

[0059] In certain aspects, the invention provides compositions comprising a nanoparticle which comprises any one of the envelopes of the invention.

[0060] In certain embodiments, the nanoparticle is ferritin self assembling nanoparticle.

[0061] In certain aspects, the invention provides a method of inducing an immune response in a subject comprising administering an immunogenic composition comprising any one of the stabilized envelopes of the invention. In certain embodiments, the composition is administered as a prime and/or a boost. In certain embodiments, the composition comprises nanoparticles. In certain embodiments, methods of the invention further comprise administering an adjuvant.

[0062] In certain aspects, the invention provides a composition comprising a plurality of nanoparticles comprising a plurality of the envelopes/trimers of the invention. In non-limiting embodiments, the envelopes/trimers of the invention are multimeric when comprised in a nanoparticle. The nanoparticle size is suitable for delivery. In non-limiting embodiments the nanoparticles are ferritin based nanoparticles.

[0063] In certain aspects, the invention provides nucleic acids comprising sequences encoding polypeptides of the invention. In certain embodiments, the nucleic acids are DNAs. In certain embodiments, the nucleic acids are mRNAs. In certain aspects, the invention provides expression vectors comprising the nucleic acids of the invention.

[0064] In certain aspects, the invention provides a pharmaceutical composition comprising mRNAs encoding the inventive envelopes. In certain embodiments, these are optionally formulated in lipid nanoparticles (LNPs). In certain embodiments, the mRNAs are modified. Modifications include without limitations modified ribonucleotides, poly-A tail, 5' cap.

[0065] In certain aspects the invention provides nucleic acids encoding the inventive polypeptide designs. In non-limiting embodiments, the nucleic acids are mRNA, modified or unmodified, suitable for use any use, e.g but not limited to use as pharmaceutical compositions. In certain embodiments, the nucleic acids are formulated in lipid, such as but not limited to LNPs.

BRIEF DESCRIPTION OF THE DRAWINGS

[0066] The patent or application file contains at least one drawing executed in color. To conform to the requirements for PCT patent applications, many of the figures presented herein are black and white representations of images originally created in color.

[0067] **Figure 1A** shows schematic of signature based approach of immunogen design. See also Bricault et al. Cell Host Microbe 2019 25(1) 59-72. **Figure 1B** shows CH505 V2 Mature Optimized Design. Shown are CH505 amino acid substitutions that are statistically associated with for V2 apex mature bNAb sensitivity. The letters represent single amino acids, and the height of the letter in the sequence LOGO indicates its frequency in the population. The numbers underneath the LOGO are HXB2 reference strain positions in the viral sequence. O stands for an N embedded in a N-linked glycosylation site. Blue are amino acids that are associated with sensitivity, red are amino acids associated with resistance, black are amino acids that were not associated with either sensitivity or resistance. The V2 SET OPT chimeric SOSIP (last row) carries all the design mutations from the full length CH505 TF V2 SET OPT except at 31, 33 and 588, 644. For the former, the SOSIP construct has the favorable mutations.

[0068] For 588, we suggest mutating to K (quite common aa, signature p-value = 0.0005-0.026 depending on the V2 bnab, odd's ratio (OR) = 2-5). For 644, we suggest mutating to R (most common sensitive aa, p=0.0006-0.007, OR=2.7-7.9).

[0069] To minimize the number of constructs, we propose adding these to UCA OPT1 SOSIP constructs (note: our UCA OPT1 carried all the sensitive signatures for mature bNAbs also in addition to most UCAs/intermediates). Gp41 mutations could also be added in some embodiments. The CH505 V2 Mature OPT design has 11 mutations as shown at positions

130, 161, 169, 170, 172, 316, 332, 335, 632, 644, and 717. These replaces 8 resistant or non-significant with sensitive signatures; One sensitive to more sensitive (R169K), one neutral to neutral (E170Q, remove charge), one resistance signature for completing glycan shield (NxST332; no impact on sensitivity).

[0070] **Figure 2** shows additional signature amino acids associated with V2 bNAb unmutated common ancestor or early intermediate antibodies from early stages of V2 apex bNAb maturation. See figure 1 for details. UCA OPT1 SOSIP construct just has one sub-optimal aa at PG9 germline reverted Ab signature sites as compared to the full length UCA OPT1 – it has an M-535 instead of I-535. We suggest using I-535 (fairly common aa, signature $p = 0.01$, OR = 3.3). Data not shown for other V2 UCAs/intermediates (CH04, PCT64) but the SOSIP UCA OPT1 construct carries all the favorable mutations for their signature sites as well. Gp41 mutations could also be added in some embodiments. The CH505 TF UCA OPT1 design includes mature V2 apex signatures, with 5 additional changes as shown at positions 161, 200, 305, 322, and 732 for UCAs.

[0071] **Figures 3A-3C** show non-limiting embodiments of amino acid sequences. Figure 3A shows amino acid sequences of envelopes. These are continuous sequences where dashes represents gaps if these sequences were aligned.

[0072] **Figures 4A and 4B** show non-limiting embodiments of amino acid and nucleic acid sequences. In Figure 4B, VDAT = cloning site and Kozak sequence. Underlined = signal peptide that is cleaved from mature protein.

[0073] **Figures 5A, 5B, 5C and 5D** show non-limiting embodiment of sortase designs and nucleic acid and protein sequences. **Figure 5E** shows non-limiting embodiments of ferritin designs. The linker between the envelope sequence and the ferritin protein sequence could be any suitable linker. The ferritin protein could be any suitable ferritin. See e.g. without limitation WO/2018/005558. The envelopes in these designs are CH505 T/F or CH505 M5. A skilled artisan can readily incorporate the V2 optimization into these envelopes.

[0074] **Figure 6** shows analyses for CAP256 IA4. For CAP256 IA4 weak signatures found due to low statistical power (3 out of 208 viruses neutralized). Only resistant signatures outside the epitope. Change to neutral at most sites would involve mutation to rare amino acid and/or removing glycans that could introduce vulnerable gaps in the glycan shield. Only two mutations introduce at 736 & 842. Designed UCA optimized constructs without (UCA OPT1) and with (UCA OPT2) these weak signatures.

[0075] **Figure 7** shows Hypervariable Loop Characteristic. Hypervariable loops cannot be aligned due to extreme length & sequence variation. Tested for associations with net charge, length & number of glycans. Found two significant hypervariable loop associations with sensitivity to V2 apex bNAbs: Positively charged V2 loops; V2 apex bNAbs have long anionic CDRH3. Smaller hypervariable V1 & V2 combined: possible steric hindrance due to the dynamic loops.

[0076] **Figures 8A-B** shows Hypervariable V1 & V2 substitutions: Optimizing for Positive Charge (**Figure 8A**) and optimizing for smaller length based on M-group Hypervariable length distribution (**Figure 8B**).

[0077] **Figures 9A-9E** show that mature signature introduction increases sensitivity to neutralization by mature V2 bNAbs. Shown are results for CH505 TF and CH505 V2 SET envelopes as gp160 constructs in a pseudovirus neutralization assay. The assay is a standard TZM-B1 cell neutralization assay as described in Sarzotti-Kelsoe et al. J Immunol Methods. 2014 Jul;409:131-46. doi: 10.1016/j.jim.2013.11.022. Epub 2013 Dec 1. Antibody is shown in each panel.

[0078] **Figures 10A-10E** show that germline signatures further increase sensitivity to neutralization by mature V2 bNAbs. Shown are results for CH505 TF, CH505 V2 SET, and CH505 UCA OPT1 envelopes as gp160 constructs in a pseudovirus neutralization assay. Antibody is shown in each panel. The thick arrow shows CH505 UCA OPT1 curve, which in panels A and E overlaps with CH505 V2 SET curve.

[0079] **Figures 11A-11E** show that UCA signatures increase neutralization sensitivity of CH505 envelopes by unmutated common ancestor (UCA) or reverted common ancestor (RUA) antibodies. Shown are results for CH505 TF, CH505 V2 SET, and CH505 UCA OPT1 envelopes as gp160 constructs in a pseudovirus neutralization assay. Antibody is shown in each panel. UCA signatures increased the sensitivity of CH505 to neutralization by both CH01 and the PCT64 V2 bNAbs UCAs. V2 SET OPT also gains CH01 UCA sensitivity, likely due to H-130. UCA OPT2 that had CAP256 VRC26 UCA signatures did not confer sensitivity to this UCA.

[0080] **Figure 12** shows a summary of the neutralization data. The table shows that introduction of V2 apex mature signatures in CH505 TF improved sensitivity to mature bNAbs, and gained sensitivity to CH01 UCA—SET OPT column. Introduction of UCA signatures further improved sensitivity to mature bNAbs, to CH01 UCA and gained sensitivity to PCT64 LMCA—UCA OPT column. In this figure the UCA OPT label shows

UCA OPT2 + N332---see the slope of the curve in Figure 15E, where the curve for CH505 UCA OPT2 + N332 is bending for the PCT64LMCA, whereas it is not for PG9RUA. This indicates that when measured the neutralization up to 250ug/ml, 50% neutralization could be reached at 105ug/ml. First column lists the antibody. “WT” refers to CH505 TF sequences without optimization signatures.

[0081] **Figures 13A-13B** shows summary of expression and binding data for various optimized designs expressed as SOSIP designs. Various non-limiting embodiments of SOSIP designs are shown in Figures 3 and 4. SET OPT & UCA OPT constructs expressed as chimeric CH505-BG505 SOSIPs. Different constructs tested with varying quality & expression. Expression of UCA OPT1 with NxST 332 and gp41 mutations resulted in highest level of trimer formation (88% versus 12% monomer) as shown in Fig 13A. Figure 13B shows antibody binding consistent with neutralization results. Binding data consistent with neutralization results.

[0082] **Figures 14A-14F** show sensitivity to neutralization by mature V2 apex bnAbs. Respective antibodies are listed in each panel. N332 represents a predicted V2 apex bNab resistance signature, but is critical for V3 bNabs (CH505 Env has N334). Moving the N334 glycan to N332 did not reduce its sensitivity to mature V2 bNabs, and rendered it highly sensitive to PGT121. The legend listed in Figure 14A is applicable to all panels in this figure.

[0083] **Figures 15A-15E** show sensitivity of UCA signatures optimized CH505 envelopes of to UCA antibodies neutralization. Respective antibodies are listed in each panel. UCA signatures increased the sensitivity of CH505 to neutralization by both CH01 and the PCT64 V2 bNab UCAs. The legend listed in Figure 14A is applicable to all panels in this figure.

DETAILED DESCRIPTION OF THE INVENTION

[0084] The development of a safe, highly efficacious prophylactic HIV-1 vaccine is of paramount importance for the control and prevention of HIV-1 infection. A major goal of HIV-1 vaccine development is the induction of broadly neutralizing antibodies (bnAbs) (Immunol. Rev. 254: 225-244, 2013). BnAbs are protective in rhesus macaques against SHIV challenge, but as yet, are not induced by current vaccines.

[0085] For the past 25 years, the HIV vaccine development field has used single or prime boost heterologous Envs as immunogens, but to date has not found a regimen to induce high levels of bnAbs.

[0086] Recently, a new paradigm for design of strategies for induction of broadly neutralizing antibodies was introduced, that of B cell lineage immunogen design (Nature Biotech. 30: 423, 2012) in which the induction of bnAb lineages is recreated. It was recently demonstrated the power of mapping the co-evolution of bnAbs and founder virus for elucidating the Env evolution pathways that lead to bnAb induction (Nature 496: 469, 2013).

[0087] Sequences/Clones

[0088] Described herein are nucleic and amino acids sequences of HIV-1 envelopes. The sequences for use as immunogens are in any suitable form. In certain embodiments, the described HIV-1 envelope sequences are gp160s. In certain embodiments, the described HIV-1 envelope sequences are gp120s. Other sequences, for example but not limited to stable SOSIP trimer designs, gp145s, gp140s, both cleaved and uncleaved, gp140 Envs with the deletion of the cleavage (C) site, fusion (F) and immunodominant (I) region in gp41-- named as gp140 Δ CFI (gp140CFI), gp140 Envs with the deletion of only the cleavage (C) site and fusion (F) domain -- named as gp140 Δ CF (gp140CF), gp140 Envs with the deletion of only the cleavage (C)---named gp140 Δ C (gp140C) (See e.g. Liao et al. Virology 2006, 353, 268-282), gp150s, gp41s, which are readily derived from the nucleic acid and amino acid gp160 sequences. In certain embodiments the nucleic acid sequences are codon optimized for optimal expression in a host cell, for example a mammalian cell, a rBCG cell or any other suitable expression system.

[0089] An HIV-1 envelope has various structurally defined fragments/forms: gp160; gp140-- including cleaved gp140 and uncleaved gp140 (gp140C), gp140CF, or gp140CFI, gp120 and gp41. A skilled artisan appreciates that these fragments/forms are defined not necessarily by their crystal structure, but by their design and bounds within the full length of the gp160 envelope. While the specific consecutive amino acid sequences of envelopes from different strains are different, the bounds and design of these forms are well known and characterized in the art.

[0090] For example, it is well known in the art that during its transport to the cell surface, the gp160 polypeptide is processed and proteolytically cleaved to gp120 and gp41 proteins. Cleavages of gp160 to gp120 and gp41 occurs at a conserved cleavage site "REKR." See Chakrabarti et al. Journal of Virology vol. 76, pp. 5357-5368 (2002) see for example Figure 1, and Second paragraph in the Introduction on p. 5357; Binley et al. Journal of Virology vol. 76, pp. 2606-2616 (2002) for example at Abstract; Gao et al. Journal of Virology vol. 79, pp. 1154-1163 (2005); Liao et al. Virology vol. 353(2): 268-282 (2006).

[0091] The role of the furin cleavage site was well understood both in terms of improving cleave efficiency, *see* Binley et al. *supra*, and eliminating cleavage, *see* Bosch and Pawlita, *Virology* 64 (5):2337-2344 (1990); Guo et al. *Virology* 174: 217-224 (1990); McCune et al. *Cell* 53:55-67 (1988); Liao et al. *J Virol.* Apr;87(8):4185-201 (2013).

[0092] Likewise, the design of gp140 envelope forms is also well known in the art, along with the various specific changes which give rise to the gp140C (uncleaved envelope), gp140CF and gp140CFI forms. Envelope gp140 forms are designed by introducing a stop codon within the gp41 sequence. *See* Chakrabarti et al. at Figure 1.

[0093] Envelope gp140C refers to a gp140 HIV-1 envelope design with a functional deletion of the cleavage (C) site, so that the gp140 envelope is not cleaved at the furin cleavage site. The specification describes cleaved and uncleaved forms, and various furin cleavage site modifications that prevent envelope cleavage are known in the art. In some embodiments of the gp140C form, two of the R residues in and near the furin cleavage site are changed to E, e.g., **RRVVEREKR** is changed to **ERVVEREKE**, and is one example of an uncleaved gp140 form. Another example is the gp140C form which has the REKR site changed to SEKS. *See supra* for references.

[0094] Envelope gp140CF refers to a gp140 HIV-1 envelope design with a deletion of the cleavage (C) site and fusion (F) region. Envelope gp140CFI refers to a gp140 HIV-1 envelope design with a deletion of the cleavage (C) site, fusion (F) and immunodominant (I) region in gp41. *See* Chakrabarti et al. *Journal of Virology* vol. 76, pp. 5357-5368 (2002) *see* for example Figure 1, and Second paragraph in the Introduction on p. 5357; Binley et al. *Journal of Virology* vol. 76, pp. 2606-2616 (2002) for example at Abstract; Gao et al. *Journal of Virology* vol. 79, pp. 1154-1163 (2005); Liao et al. *Virology* vol. 353(2): 268-282 (2006).

[0095] In certain embodiments, the envelope design in accordance with the present invention involves deletion of residues (e.g., 5-11, 5, 6, 7, 8, 9, 10, or 11 amino acids) at the N-terminus. For delta N-terminal design, amino acid residues ranging from 4 residues or even fewer to 14 residues or even more are deleted. These residues are between the maturation (signal peptide, usually ending with CX, X can be any amino acid) and "VPVXXXX...". In case of CH505 T/F Env as an example, 8 amino acids (*italicized and underlined* in the below sequence) were deleted:

MRVMGIQRNYPQWWIWSMLGFWMLMICNGM*WVTVYYG*VPVWKEAKTTLFCASDA
KAYEKEVHNVWATHACVPTDPNPQE...(rest of envelope sequence is indicated as "...").

In other embodiments, the delta N-design described for CH505 T/F envelope can be used to

make delta N-designs of other CH505 envelopes. In certain embodiments, the invention relates generally to an immunogen, gp160, gp120 or gp140, without an N-terminal Herpes Simplex gD tag substituted for amino acids of the N-terminus of gp120, with an HIV leader sequence (or other leader sequence), and without the original about 4 to about 25, for example 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 amino acids of the N-terminus of the envelope (e.g. gp120). See WO2013/006688, e.g. at pages 10-12, the contents of which publication is hereby incorporated by reference in its entirety.

[0096] The general strategy of deletion of N-terminal amino acids of envelopes results in proteins, for example gp120s, expressed in mammalian cells that are primarily monomeric, as opposed to dimeric, and, therefore, solves the production and scalability problem of commercial gp120 Env vaccine production. In other embodiments, the amino acid deletions at the N-terminus result in increased immunogenicity of the envelopes.

[0097] In certain aspects, the invention provides composition and methods which CH505 Envs, as gp120s, gp140s cleaved and uncleaved, gp145s, gp150s and gp160s, stabilized and/or multimerized trimers, as proteins, DNAs, RNAs, or any combination thereof, administered as primes and boosts to elicit immune response. CH505 Envs as proteins would be co-administered with nucleic acid vectors containing Envs to amplify antibody induction. In certain embodiments, the compositions and methods include any immunogenic HIV-1 sequences to give the best coverage for T cell help and cytotoxic T cell induction. In certain embodiments, the compositions and methods include mosaic and/or consensus HIV-1 genes to give the best coverage for T cell help and cytotoxic T cell induction. In certain embodiments, the compositions and methods include mosaic group M and/or consensus genes to give the best coverage for T cell help and cytotoxic T cell induction. In some embodiments, the mosaic genes are any suitable gene from the HIV-1 genome. In some embodiments, the mosaic genes are Env genes, Gag genes, Pol genes, Nef genes, or any combination thereof. See e.g. US Patent No. 7951377. In some embodiments the mosaic genes are bivalent mosaics. In some embodiments the mosaic genes are trivalent. In some embodiments, the mosaic genes are administered in a suitable vector with each immunization with Env gene inserts in a suitable vector and/or as a protein. In some embodiments, the mosaic genes, for example as bivalent mosaic Gag group M consensus genes, are administered in a suitable vector, for example but not limited to HSV2, would be administered with each immunization with Env gene inserts in a suitable vector, for example but not limited to HSV-2.

[0098] **Nucleic acid sequences**

[0099] In certain aspects the invention provides compositions and methods of Env genetic immunization either alone or with Env proteins to recreate the swarms of evolved viruses that have led to bnAb induction. Nucleotide-based vaccines offer a flexible vector format to immunize against virtually any protein antigen. Currently, two types of genetic vaccination are available for testing—DNAs and mRNAs.

[0100] In certain aspects the invention contemplates using immunogenic compositions wherein immunogens are delivered as DNA. See Graham BS, Enama ME, Nason MC, Gordon II, Peel SA, et al. (2013) DNA Vaccine Delivered by a Needle-Free Injection Device Improves Potency of Priming for Antibody and CD8⁺ T-Cell Responses after rAd5 Boost in a Randomized Clinical Trial. PLoS ONE 8(4): e59340, page 9. Various technologies for delivery of nucleic acids, as DNA and/or RNA, so as to elicit immune response, both T-cell and humoral responses, are known in the art and are under developments. In certain embodiments, DNA can be delivered as naked DNA. In certain embodiments, DNA is formulated for delivery by a gene gun. In certain embodiments, DNA is administered by electroporation, or by a needle-free injection technologies, for example but not limited to Biojector® device. In certain embodiments, the DNA is inserted in vectors. The DNA is delivered using a suitable vector for expression in mammalian cells. In certain embodiments the nucleic acids encoding the envelopes are optimized for expression. In certain embodiments DNA is optimized, e.g. codon optimized, for expression. In certain embodiments the nucleic acids are optimized for expression in vectors and/or in mammalian cells. In non-limiting embodiments these are bacterially derived vectors, adenovirus based vectors, rAdenovirus (e.g. Barouch DH, et al. Nature Med. 16: 319-23, 2010), recombinant mycobacteria (e.g. rBCG or M smegmatis) (Yu, JS et al. Clinical Vaccine Immunol. 14: 886-093, 2007; ibid 13: 1204-11, 2006), and recombinant vaccinia type of vectors (Santra S. Nature Med. 16: 324-8, 2010), for example but not limited to ALVAC, replicating (Kibler KV et al., PLoS One 6: e25674, 2011 nov 9.) and non-replicating (Perreau M et al. J. virology 85: 9854-62, 2011) NYVAC, modified vaccinia Ankara (MVA)), adeno-associated virus, Venezuelan equine encephalitis (VEE) replicons, Herpes Simplex Virus vectors, and other suitable vectors.

[0101] In certain aspects the invention contemplates using immunogenic compositions wherein immunogens are delivered as DNA or RNA in suitable formulations. Various technologies which contemplate using DNA or RNA or may use complexes of nucleic acid

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

[0102] In certain aspects, the invention provides nucleic acids comprising sequences encoding envelopes of the invention. In certain embodiments, the nucleic acids are DNAs. In certain embodiments, the nucleic acids are mRNAs. In certain aspects, the invention provides expression vectors comprising the nucleic acids of the invention.

[0103] In certain aspects, the invention provides a pharmaceutical composition comprising mRNAs encoding the inventive antibodies. In certain embodiments, these are optionally formulated in lipid nanoparticles (LNPs). In certain embodiments, the mRNAs are modified. Modifications include without limitations modified ribonucleotides, poly-A tail, 5' cap.

[0104] In certain aspects the invention provides nucleic acids encoding the inventive envelopes. In non-limiting embodiments, the nucleic acids are mRNA, modified or unmodified, suitable for use any use, e.g. but not limited to use as pharmaceutical compositions. In certain embodiments, the nucleic acids are formulated in lipid, such as but not limited to LNPs.

[0105] In some embodiments the antibodies are administered as nucleic acids, including but not limited to mRNAs which could be modified and/or unmodified. See US Pub 20180028645A1, US Pub 20090286852, US Pub 20130111615, US Pub 20130197068, US Pub 20130261172, US Pub 20150038558, US Pub 20160032316, US Pub 20170043037, US Pub 20170327842, US Patent 10,006,007, US Patent 9,371,511, US Patent 9,012,219, US Pub 20180265848, US Pub 20170327842, US Pub 20180344838A1 at least at paragraphs [0260] –[0281], WO/2017/182524 for non-limiting embodiments of chemical modifications, wherein each content is incorporated by reference in its entirety.

[0106] mRNAs delivered in LNP formulations have advantages over non-LNPs formulations. See US Pub 20180028645A1, WO/2018/081638, WO/2016/176330, wherein each content is incorporated by reference in its entirety.

[0107] In certain embodiments the nucleic acid encoding an envelope is operably linked to a promoter inserted in an expression vector. In certain aspects the compositions comprise a suitable carrier. In certain aspects the compositions comprise a suitable adjuvant.

[0108] In certain aspects the invention provides an expression vector comprising any of the nucleic acid sequences of the invention, wherein the nucleic acid is operably linked to a promoter. In certain aspects the invention provides an expression vector comprising a nucleic acid sequence encoding any of the polypeptides of the invention, wherein the nucleic acid is operably linked to a promoter. In certain embodiments, the nucleic acids are codon optimized for expression in a mammalian cell, in vivo or in vitro. In certain aspects the invention provides nucleic acids comprising any one of the nucleic acid sequences of invention. In certain aspects the invention provides nucleic acids consisting essentially of any one of the nucleic acid sequences of invention. In certain aspects the invention provides nucleic acids consisting of any one of the nucleic acid sequences of invention. In certain embodiments the nucleic acid of the invention, is operably linked to a promoter and is inserted in an expression vector. In certain aspects the invention provides an immunogenic composition comprising the expression vector.

[0109] In certain aspects the invention provides a composition comprising at least one of the nucleic acid sequences of the invention. In certain aspects the invention provides a composition comprising any one of the nucleic acid sequences of invention. In certain aspects the invention provides a composition comprising at least one nucleic acid sequence encoding any one of the polypeptides of the invention.

[0110] In one embodiment, the nucleic acid is an RNA molecule. In one embodiment, the RNA molecule is transcribed from a DNA sequence described herein. In some embodiments, the RNA molecule is encoded by one of the inventive sequences. In another embodiment, the nucleotide sequence comprises an RNA sequence transcribed by a DNA sequence encoding the polypeptide sequence of the sequences of the invention, or a variant thereof or a fragment thereof. Accordingly, in one embodiment, the invention provides an RNA molecule encoding one or more of inventive antibodies. The RNA may be plus-stranded. Accordingly, in some embodiments, the RNA molecule can be translated by cells without needing any intervening replication steps such as reverse transcription.

[0111] In some embodiments, a RNA molecule of the invention may have a 5' cap (e.g. but not limited to a 7-methylguanosine, 7mG(5')ppp(5')NlmpNp). This cap can enhance in vivo translation of the RNA. The 5' nucleotide of an RNA molecule useful with the invention may

have a 5' triphosphate group. In a capped RNA this may be linked to a 7-methylguanosine via a 5'-to-5' bridge. A RNA molecule may have a 3' poly-A tail. It may also include a poly-A polymerase recognition sequence (e.g. AAUAAA) near its 3' end. In some embodiments, a RNA molecule useful with the invention may be single-stranded. In some embodiments, a RNA molecule useful with the invention may comprise synthetic RNA.

[0112] The recombinant nucleic acid sequence can be an optimized nucleic acid sequence. Such optimization can increase or alter the immunogenicity of the envelope. Optimization can also improve transcription and/or translation. Optimization can include one or more of the following: low GC content leader sequence to increase transcription; mRNA stability and codon optimization; addition of a kozak sequence (e.g., GCC ACC) for increased translation; addition of an immunoglobulin (Ig) leader sequence encoding a signal peptide; and eliminating to the extent possible cis-acting sequence motifs (i.e., internal TATA boxes).

[0113] Methods for in vitro transfection of mRNA and detection of envelope expression are known in the art.

[0114] Methods for expression and immunogenicity determination of nucleic acid encoded envelopes are known in the art.

[0115] In certain aspects the invention contemplates using immunogenic compositions wherein immunogens are delivered as recombinant proteins. Various methods for production and purification of recombinant proteins, including trimers such as but not limited to SOSIP based trimers, suitable for use in immunization are known in the art. In certain embodiments recombinant proteins are produced in CHO cells.

[0116] The immunogenic envelopes can also be administered as a protein boost in combination with a variety of nucleic acid envelope primes (e.g., HIV -1 Envs delivered as DNA expressed in viral or bacterial vectors).

[0117] Dosing of proteins and nucleic acids can be readily determined by a skilled artisan. A single dose of nucleic acid can range from a few nanograms (ng) to a few micrograms (μg) or milligram of a single immunogenic nucleic acid. Recombinant protein dose can range from a few μg micrograms to a few hundred micrograms, or milligrams of a single immunogenic polypeptide.

[0118] Administration: The compositions can be formulated with appropriate carriers using known techniques to yield compositions suitable for various routes of administration. In certain embodiments the compositions are delivered via intramuscular (IM), via

subcutaneous, via intravenous, via nasal, via mucosal routes, or any other suitable route of immunization.

[0119] The compositions can be formulated with appropriate carriers and adjuvants using techniques to yield compositions suitable for immunization. The compositions can include an adjuvant, such as, for example but not limited to, alum, 3M052, poly IC, MF-59 or other squalene-based adjuvant, ASOIB, or other liposomal based adjuvant suitable for protein or nucleic acid immunization. In certain embodiments, the adjuvant is GSK AS01E adjuvant containing MPL and QS21. This adjuvant has been shown by GSK to be as potent as the similar adjuvant AS01B but to be less reactogenic using HBsAg as vaccine antigen [Leroux-Roels et al., IABS Conference, April 2013]. In certain embodiments, TLR agonists are used as adjuvants. In other embodiment, adjuvants which break immune tolerance are included in the immunogenic compositions.

[0120] In certain embodiments, the compositions and methods comprise any suitable agent or immune modulation which could modulate mechanisms of host immune tolerance and release of the induced antibodies. In non-limiting embodiments modulation includes PD-1 blockade; T regulatory cell depletion; CD40L hyperstimulation; soluble antigen administration, wherein the soluble antigen is designed such that the soluble agent eliminates B cells targeting dominant epitopes, or a combination thereof. In certain embodiments, an immunomodulatory agent is administered in at time and in an amount sufficient for transient modulation of the subject's immune response so as to induce an immune response which comprises broad neutralizing antibodies against HIV-1 envelope. Non-limiting examples of such agents is any one of the agents described herein: e.g. chloroquine (CQ), PTP1B Inhibitor - CAS 765317-72-4 - Calbiochem or MSI 1436 clodronate or any other bisphosphonate; a Foxo1 inhibitor, e.g. 344355 | Foxo1 Inhibitor, AS1842856 - Calbiochem; Gleevec, anti-CD25 antibody, anti-CCR4 Ab, an agent which binds to a B cell receptor for a dominant HIV-1 envelope epitope, or any combination thereof. In non-limiting embodiments, the modulation includes administering an anti-CTLA4 antibody. Non-limiting examples are ipilimumab and tremelimumab. In certain embodiments, the methods comprise administering a second immunomodulatory agent, wherein the second and first immunomodulatory agents are different.

[0121] There are various host mechanisms that control bnAbs. For example, highly somatically mutated antibodies become autoreactive and/or less fit (Immunity 8: 751, 1998; PloS Comp. Biol. 6 e1000800, 2010; J. Thoret. Biol. 164:37, 1993); Polyreactive/autoreactive

naïve B cell receptors (unmutated common ancestors of clonal lineages) can lead to deletion of Ab precursors (Nature 373: 252, 1995; PNAS 107: 181, 2010; J. Immunol. 187: 3785, 2011); Abs with long HCDR3 can be limited by tolerance deletion (JI 162: 6060, 1999; JCI 108: 879, 2001). BnAb knock-in mouse models are providing insights into the various mechanisms of tolerance control of MPER BnAb induction (deletion, anergy, receptor editing). Other variations of tolerance control likely will be operative in limiting BnAbs with long HCDR3s, high levels of somatic hypermutations.

[0122] For a summary of CH505 sequences and designs see WO2017151801, e.g. but not limited to Table 1, Figures 22-24, and WO2014042669 (Figure 17).

[0123] It is readily understood that the envelope glycoproteins referenced in various examples and figures comprise a signal/leader sequence. It is well known in the art that HIV-1 envelope glycoprotein is a secretory protein with a signal or leader peptide sequence that is removed during processing and recombinant expression (without removal of the signal peptide, the protein is not secreted). *See* for example Li et al. Control of expression, glycosylation, and secretion of HIV-1 gp120 by homologous and heterologous signal sequences. Virology 204(1):266-78 (1994) ("Li et al. 1994"), at first paragraph, and Li et al. Effects of inefficient cleavage of the signal sequence of HIV-1 gp120 on its association with calnexin, folding, and intracellular transport. PNAS 93:9606-9611 (1996) ("Li et al. 1996"), at 9609, Any suitable signal sequence could be used. In some embodiments the leader sequence is the endogenous leader sequence. Most of the gp120 and gp160 amino acid sequences include the endogenous leader sequence. In other non-limiting examples, the leader sequence is human Tissue Plasminogen Activator (TPA) sequence, human CD5 leader sequence (e.g. MPMGSLQPLATLYLLGMLVASVLA). Most of the chimeric designs include CD5 leader sequence. A skilled artisan appreciates that when used as immunogens, and for example when recombinantly produced, the amino acid sequences of these proteins do not comprise the leader peptide sequences.

[0124] HIV-1 envelope trimers and other envelope designs

[0125] This example shows that stabilized HIV-1 Env trimer immunogens show enhanced antigenicity for broadly neutralizing antibodies and are not recognized by non-neutralizing antibodies. The example also describes additional envelope modifications and designs. In some embodiments these envelopes, including but not limited to trimers are further multimerized, and/or used as particulate, high-density array in liposomes or other particles,

for example but not limited to nanoparticles. Any one of the envelopes of the invention could be designed and expressed as described herein.

[0126] A stabilized chimeric SOSIP designs were used to generate CH505 trimers. This design was applicable to diverse viruses from multiple clades.

[0127] Elicitation of neutralizing antibodies is one goal for antibody-based vaccines. Neutralizing antibodies target the native trimeric HIV-1 Env on the surface virions. The trimeric HIV-1 envelope protein consists of three protomers each containing a gp120 and gp41 heterodimer. Recent immunogen design efforts have generated soluble near-native mimics of the Env trimer that bind to neutralizing antibodies but not non-neutralizing antibodies. The recapitulation of the native trimer could be a key component of vaccine induction of neutralizing antibodies. Neutralizing Abs target the native trimeric HIV-1 Env on the surface of viruses (Poignard et al. *J Virol.* 2003 Jan;77(1):353-65; Parren et al. *J Virol.* 1998 Dec;72(12):10270-4.; Yang et al. *J Virol.* 2006 Nov;80(22):11404-8.). The HIV-1 Env protein consists of three protomers of gp120 and gp41 heterodimers that are noncovalently linked together (Center et al. *J Virol.* 2002 Aug;76(15):7863-7.). Soluble near-native trimers preferentially bind neutralizing antibodies as opposed to non-neutralizing antibodies (Sanders et al. *PLoS Pathog.* 2013 Sep; 9(9): e1003618).

[0128] Sequential Env vaccination has elicited broad neutralization in the plasma of one macaque. The overall goal of our project is to increase the frequency of vaccine induction of bnabs in the plasma of primates with Env vaccination. We hypothesized that vaccination with immunogens that target bnAb B cell lineage and mimic native trimers will increase the frequency of broadly neutralizing plasma antibodies. One goal is increasing the frequency of vaccine induction of bnAb in the plasma of primates by Env vaccination. It is expected that vaccination with immunogens that target bnAb B cell lineages and mimic the native trimers on virions will increase the frequency of broadly neutralizing plasma antibodies.

[0129] Previous work has shown that CH505 derived soluble trimers are hard to produce. From a study published by Julien et al in 2015 (*Proc Natl Acad Sci U S A.* 2015 Sep 22; 112(38): 11947–11952.) it was shown that while CH505 produced comparable amounts of protein by transient transfection, only 5% of the CH505 protein formed trimer which 5 times lower than the gold standard viral strain BG505. Provided here are non-limiting embodiments of well-folded trimers for Env immunizations.

[0130] Near-native soluble trimers using the 6R.SOSIP.664 design are capable of generating autologous tier 2 neutralizing plasma antibodies in the plasma (Sanders et al. 2015), which

provides a starting point for designing immunogens to elicit broadly neutralizing antibodies. While these trimers are preferentially antigenic for neutralizing antibodies, they still possess the ability to expose the V3 loop, which generally results in strain-specific binding and neutralizing antibodies after vaccination. Using the unliganded structure the BG505.6R.SOSIP.664 has been stabilized by adding cysteines at position 201 and 433 to constrain the conformational flexibility such that the V3 loop is maintained unexposed (Kwon et al. *Nat Struct Mol Biol.* 2015 Jul; 22(7): 522–531.).

[0131] Provided are engineered trimeric immunogens derived from multiple viruses from CH505. We generated chimeric 6R.SOSIP.664, chimeric disulfide stabilized (DS) 6R.SOSIP.664 (Kwon et al *Nat Struct Mol Biol.* 2015 Jul; 22(7): 522–531.), chimeric 6R.SOSIP.664v4.1 (DeTaeye et al. *Cell.* 2015 Dec 17;163(7):1702-15. doi: 10.1016/j.cell.2015.11.056), and chimeric 6R.SOSIP.664v4.2 (DeTaeye et al. *Cell.* 2015 Dec 17;163(7):1702-15. doi: 10.1016/j.cell.2015.11.056). The 6R.SOSIP.664 is the basis for all of these designs and is made as a chimera of C.CH505 and A.BG505. The gp120 of C.CH505 was fused with the BG505 inner domain gp120 sequence within the alpha helix 5 ($\alpha 5$) to result in the chimeric protein. The chimeric gp120 is disulfide linked to the A.BG505 gp41 as outlined by Sanders et al. (*PLoS Pathog.* 2013 Sep; 9(9): e1003618). These immunogens were designed as chimeric proteins that possess the BG505 gp41 connected to the CH505 gp120, since the BG505 strain is particularly adept at forming well-folded, closed trimers. This envelope design retains the CH505 CD4 binding site that is targeted by the CH103 and CH235 broadly neutralizing antibody lineages that were isolated from CH505.

[0132] Based on the various designs, any other suitable envelope, for example but not limited to CH505 envelopes as described in WO2014042669 can be designed.

[0133] Recombinant envelopes as trimers could be produced and purified by any suitable method. For a non-limiting example of purification methods see Ringe RP, Yasmeen A, Ozorowski G, Go EP, Pritchard LK, Guttman M, Ketas TA, Cottrell CA, Wilson IA, Sanders RW, Cupo A, Crispin M, Lee KK, Desaire H, Ward AB, Klasse PJ, Moore JP. 2015. Influences on the design and purification of soluble, recombinant native-like HIV-1 envelope glycoprotein trimers. *J Virol* 89:12189 -12210. doi:10.1128/JVI.01768-15.

[0134] **Multimeric Envelopes**

[0135] Presentation of antigens as particulates reduces the B cell receptor affinity necessary for signal transduction and expansion (See Baptista et al. *EMBO J.* 2000 Feb 15; 19(4): 513–520). Displaying multiple copies of the antigen on a particle provides an avidity effect that

can overcome the low affinity between the antigen and B cell receptor. The initial B cell receptor specific for pathogens can be low affinity, which precludes vaccines from being able to stimulate and expand B cells of interest. In particular, very few naïve B cells from which HIV-1 broadly neutralizing antibodies arise can bind to soluble HIV-1 Envelope. Provided are envelopes, including but not limited to trimers as particulate, high-density array on liposomes or other particles, for example but not limited to nanoparticles. See e.g. He et al. Nature Communications 7, Article number: 12041 (2016), doi:10.1038/ncomms12041; Bamrungsap et al. Nanomedicine, 2012, 7 (8), 1253-1271.

[0136] To improve the interaction between the naïve B cell receptor and immunogens, envelope designed can be created to wherein the envelope is presented on particles, e.g. but not limited to nanoparticle. In some embodiments, the HIV-1 Envelope trimer could be fused to ferritin. Ferritin protein self assembles into a small nanoparticle with three fold axis of symmetry. At these axes the envelope protein is fused. Therefore, the assembly of the three-fold axis also clusters three HIV-1 envelope protomers together to form an envelope trimer. Each ferritin particle has 8 axes which equates to 8 trimers being displayed per particle. See e.g. Sliepen et al. Retrovirology 2015 12:82, DOI: 10.1186/s12977-015-0210-4.

[0137] Any suitable ferritin sequence could be used. In non-limiting embodiments, ferritin sequences are disclosed in WO/2018/005558.

[0138] Ferritin nanoparticle linkers: The ability to form HIV-1 envelope ferritin nanoparticles relies self-assembly of 24 ferritin subunits into a single ferritin nanoparticle. The addition of a ferritin subunit to the C-terminus of HIV-1 envelope may interfere with the ability of the ferritin subunit to fold properly and or associate with other ferritin subunits. When expressed alone ferritin readily forms 24-subunit nanoparticles, however appending it to envelope only yields nanoparticles for certain envelopes. Since the ferritin nanoparticle forms in the absence of envelope, the envelope could be sterically hindering the association of ferritin subunits. Thus, we designed ferritin with elongated glycine-serine linkers to further distance the envelope from the ferritin subunit. To make sure that the glycine linker is attached to ferritin at the correct position, we created constructs that attach at second amino acid position or the fifth amino acid position. The first four n-terminal amino acids of natural *Helicobacter pylori* ferritin are not needed for nanoparticle formation but may be critical for proper folding and oligomerization when appended to envelope. Thus, we designed constructs with and without the Leucine, serine, and lysine amino acids following the glycine-serine linker. The goal will be to find a linker length that is suitable for formation of envelope nanoparticles when ferritin

is appended to most envelopes. Any suitable linker between the envelope and ferritin could be used, so long as the fusion protein is expressed and the trimer is formed.

[0139] Another approach to multimerize expression constructs uses staphylococcus Sortase A transpeptidase ligation to conjugate inventive envelope trimers, for e.g. but not limited to cholesterol. Non-limiting embodiments of envelope designs for use in Sortase A reaction are shown in Figures 5A-B. The trimers can then be embedded into liposomes via the conjugated cholesterol. To conjugate the trimer to cholesterol either a C-terminal LPXTG tag or a N-terminal pentaglycine repeat tag is added to the envelope trimer gene. Cholesterol is also synthesized with these two tags. Sortase A is then used to covalently bond the tagged envelope to the cholesterol. The sortase A-tagged trimer protein can also be used to conjugate the trimer to other peptides, proteins, or fluorescent labels. In non-limiting embodiments, the sortase A tagged trimers are conjugated to ferritin to form nanoparticles.

[0140] The invention provides design of envelopes and trimer designs wherein the envelope comprises a linker which permits addition of a lipid, such as but not limited to cholesterol, via a Sortase A reaction. See e.g. Tsukiji, S. and Nagamune, T. (2009), Sortase-Mediated Ligation: A Gift from Gram-Positive Bacteria to Protein Engineering. *ChemBioChem*, 10: 787–798. doi:10.1002/cbic.200800724; Proft, T. Sortase-mediated protein ligation: an emerging biotechnology tool for protein modification and immobilisation. *Biotechnol Lett* (2010) 32: 1. doi:10.1007/s10529-009-0116-0; Lena Schmohl, Dirk Schwarzer, Sortase-mediated ligations for the site-specific modification of proteins, *Current Opinion in Chemical Biology*, Volume 22, October 2014, Pages 122-128, ISSN 1367-5931, dx.doi.org/10.1016/j.cbpa.2014.09.020; Tabata et al. *Anticancer Res.* 2015 Aug;35(8):4411-7; Pritz et al. *J. Org. Chem.* 2007, 72, 3909-3912.

[0141] The lipid modified envelopes and trimers could be formulated as liposomes. Any suitable liposome composition is contemplated.

[0142] The lipid modified and multimerized envelopes and trimers could be formulated as liposomes. Any suitable liposome composition is contemplated.

[0143] Nomenclature for trimers: chim.6R.DS.SOSIP.664 is SOSIP.I; CHIM.6R.SOSIP.664 is SOSIP.II; CHIM.6R.SOSIP.664V4.1 is SOSIP.III.

[0144] **V2 optimization**

[0145] V2 apex bNAbs are an attractive target for immunogen design. V2 apex bNAbs arise frequently in HIV-1 infected humans (12-15%) and in SHIV infected RMs (11%). V2 apex bNAbs have low levels of somatic hypermutation are required (Wiehe et al *Cell Host*

Microbe 23(6):759 (2018)). V2 apex bNAbs have low levels of poly- and autoreactivity (Liu et al J Virol 89:784 (2015)). V2 apex bNAbs have long anionic CDRH3s (>24aa) encoded by germline. However, V2 apex bNAbs precursors are rare --germline targeting immunogens are critical. Also, no natural Envs that can target multiple V2 apex bNAb lineages.

Described herein is immunogen design to design such immunogens.

[0146] The CH505 HIV-1 virus has been subject to intensive study as a vaccine reagent based on the observation that during the course of the natural CH505 HIV-1 infection, potent broadly neutralizing antibodies were generated by the host that targeted the CD4bs region. Here we have designed an immunogen based on the surprising finding that the HIV-1 CH505 transmitted-founder (TF) virus Envelopes, when used as vaccine, have the capacity to induce V2 apex directed heterologous neutralizing antibody responses. This has been observed in a knock-in mice, rabbits and rhesus macaques, and in one CH505 SHIV infected macaque. These results raise the prospect of ultimately creating a dual-targeting CH505-based immunogen design that can induce both V2 apex and CD4bs broadly neutralizing antibodies (bNAbs). The designs we propose focuses on enhancing both the initiation of appropriate V2 apex targeting neutralizing antibody and expand the breadth of the response.

[0147] Despite the fact the CH505 TF Envelope can elicit V2 apex neutralizing antibody responses, it is not particularly sensitive to mature V2 apex bNAbs and is not neutralized by putative V2 apex bNAb precursors. We hypothesized that these factors could be limit the successful V2 apex bNAb induction, and that CH505 TF variants with improved sensitivity to V2 apex mature and precursor antibodies might serve as better immunogens.

[0148] Thus, we used our previously published statistically robust and phylogenetically corrected strategy to compare the CH505 TF to amino acid and glycan signatures that associate with sensitivity to multiple V2 apex bNAbs (*Bricault et al. Cell Host Microbe (2019) 25:59-72*). We found that CH505 TF carried resistance signatures at 10 sites, and by introducing favorable mutations at these sites, we designed a variant called V2 SET OPT (signature-based epitope targeted optimized) (Fig. 1). Shorter and more positively charged hypervariable V1 and V2 loops are significantly associated with neutralization sensitivity by mature V2 apex bNAbs, so we also introduced optimal V1 and V2 hypervariable loops from two natural Envs, ZM233.6 and T250-4, respectively, into our constructs.

[0149] For V2 Apex UCA analyses, neutralization data for 208 global viruses against CH04 & CAP256 UCAs, and heavy and/or light chain germline reverted PG9. See Gorman et al. NSMB 23 81-90 (2016). Unlike other bNAb classes, V2 apex precursors can neutralize

heterologous strains. CH04 UCA shows 4% breadth. PG9 both heavy & light chain reverted = 2% breadth. CAP256 UCA only neutralizes 1 autologous virus. Partial germline reverted PG9 (heavy or light) have higher breadth. These data were used these data to calculate signatures. The terms UCA and germline are used interchangeably.

[0150] We next applied signature analyses to neutralization data for 109-208 global viruses tested against unmutated or early ancestral antibodies that ultimately gave rise to antibody lineages that targeted the V2 apex and potent broadly neutralizing antibodies: CH04 UCA, CAP256-VRC26 and PCT64 early intermediates, and heavy and/or light chain germline reverted PG9 and PGT145. Using this strategy, we identified signatures associated with sensitivity to V2 apex precursors (Fig. 2). CH505 TF UCA OPT 1 includes mature V2 apex signatures, with 5 additional for UCAs—additional positions are 161, 200, 305, 322, and 732.

[0151] The hypervariable loop characteristics associated with sensitivity to V2 apex precursors were similar to those of the mature, and hence, the hypervariable V1 and V2 loop modifications from V2 SET OPT were retained.

[0152] In non-limiting embodiments, these vaccines are being expressed as chimeric SOSIP proteins, and so have CH505 TF gp120s, with a BG505 gp41 that ends at HIV-1 HXB2 numbering position 664. SOSIP proteins are modified Env proteins that are stabilized for expression as native-like soluble trimers.

[0153] These sensitivity mutations in a CH505 TF background expressed as SOSIP proteins we propose will result in immunogens that are more susceptible to V2-apex antibodies, and thus may be better able to trigger and stimulate them.

[0154] The modified sequence we are suggesting trying as immunogens are enclosed. We start the alignment with CH505.TF as a reference, the natural transmitted founder virus that we are building mutations into. We follow with full length protein sequences that contain the amino acid modifications we believe may be advantageous. We include the natural strains ZM233.6 and T250-4 in the alignment, as we included their hypervariable regions.

[0155] Table 1 shows V2 Optimized CH505 TF immunogens

Gene number	Protein name	Immunogen criteria
HV1301908	CH505TF_V2.UCA.OPT1.gp41mut_ch.SOSIP.v4.1	Optimized gp120 and gp41 based on V2-glycan

		bnAb UCA neutralization
HV1301909	CH505TF_V2.UCA.OPT1.N332.gp41mut_ch.SOSIP.v4.1	Optimized gp120 and gp41 based on V2-glycan bnAb UCA neutralization
HV1301910	CH505TF_V2.SET.OPT_ch.SOSIPv4.1	Optimized gp120 based on V2- glycan bnAb neutralization
HV1301911	CH505TF_V2.SET.OPT.N332_ch.SOSIPv4.1	Optimized gp120 based on V2- glycan bnAb neutralization with N332 glycan hole filled
HV1301912	CH505TF_V2.UCA.OPT1_ch.SOSIPv4.1	Optimized gp120 based on V2- glycan bnAb UCA neutralization
HV1301913	CH505TF_V2.UCA.OPT1.N332_ch.SOSIP.v4.1	Optimized gp120 based on V2- glycan bnAb UCA neutralization with N332 glycan hole filled

[0156] Non-limiting embodiments of sequences of the envelopes in Table 1 are described in Figures 3A-C and 4A-B. Figures 5A-E shows non-limiting embodiments of multimerization designs, including ferritin and/or sorstase, which could be used as guidance to design V2OPT CH505T/F designs. Non-limiting designs of envelope protomers include SOSIP designs, designs comprising F14 mutations (See WO/2020/072169), and so forth.

[0157] Throughout the application amino acid positions numbers refer to HXB2 numbering.

[0158] The invention is described in the following non-limiting examples.

EXAMPLES

Example 1

[0159] Saunders et al. have reported that vaccination with stabilized CH505 SOSIP trimers elicits V1V2-glycan bnAbs. See Cell Rep. 2017 Dec 26; 21(13): 3681–3690, incorporated by reference in its entirety.

Example 2:

[0160] CH505-BG505 Chimeric SOSIP Redesign for V2 UCA Constructs & for V5 glycan mutants

[0161] Chimeric v4 6R SOSIP constructs have BG505 gp41 and end at HXB2 664. Thus, the SOSIP constructs have sub-optimal amino acids at some of our mature and UCA signature sites in gp41.

[0162] Since the region encompassed by the SOSIP constructs ends at 664, the UCA OPT1 SOSIP and OPT2 SOSIP constructs are the same. Same for UCA OPT1 N332 and UCA OPT2 N332 SOSIPs. So, skip testing the OPT2 SOSIP constructs.

[0163] Instead, we suggest testing two other constructs: with and without gp41 optimized mutations in the backbones of UCA OPT1 and UCA OPT1 N332 – these are UCA OPT1 gp41mut and UCA OPT1 N332 gp41mut.

[0164] The gp41mut constructs introduce favorable amino acids at 3 sites: 588 and 644 (signature sites for mature V2 apex bNAbs) and 535 (PG9 germline reverted signature).

[0165] List of SOSIP constructs for testing:

[0166] CH505TF_V2.SET.OPT_ch.SOSIPv4.1

[0167] CH505TF_V2.SET.OPT.N332_ch.SOSIPv4.1

[0168] CH505TF_V2.UCA.OPT1_ch.SOSIPv4.1

[0169] CH505TF_V2.UCA.OPT1.N332_ch.SOSIP.v4.1

[0170] But we propose testing the following two instead of the UCA OPT2 constructs (since they are same as UCA OPT1 for the SOSIP constructs):

[0171] CH505TF_V2.UCA.OPT1.gp41mut_ch.SOSIP.v4.1

[0172] CH505TF_V2.UCA.OPT1.N332.gp41mut_ch.SOSIP.v4.1

[0173] The gp41mut constructs have 3 mutations in gp41: R->K at position 588; G->R at position 644; M->I at position 535.

[0174] Signatures are amino acids or glycan motifs statistically associated with one group of viruses vs others. We previously identified sequence patterns associated with sensitivity to mature V2 bNAbs. See Bricault et al. Cell Host Microbe 2019 25(1) 59-72. These analyses were used for designing CH505 OPT sequences and consider phylogenetic and/or contact site, and robust across bNAbs and datasets.

[0175] Analyses and characterization of the optimized designs are shown in Figures 1, 2, 6-13.

Example 3 Animal Studies

[0176] In non-limiting embodiment these immunogens can be used as either single primes and boosts in humanized mice or bnAb UCA or intermediate antibody VH + VL knockin mice, non-human primates (NHPs) or humans, or used in combinations in animal models or in humans.

[0177] Immunogens to initiate V1V2, and/or CD4 binding site and/or Fusion Peptide unmutated common ancestor (UCA) broadly neutralizing antibody (bnAbs) precursors.

[0178] Non-limiting examples of immunizations are listed:

1. Prime X 3 with either A, B, C, D, G or H (listed in Figures 3A-C, Table 1). In other embodiments, these immunogens could be in any suitable envelope form.
2. Take the optimal prime for bnAbs and after priming, boost with A, B, C, D, G or H.
3. Take the optimal prime for bnAbs, and after priming boost with a mixture of A, B, C, D, G or H.
4. Prime X3 with the mixture of A, B, C, D, G and H and the boost with one of A, B, C, D, D or H to focus the response on bnAb epitopes.
5. Prime as in steps #1-4 above and then boost with the CH505 Transmitted/Founder (TF) gp140 SOSIP trimer that has induced autologous neutralizing antibodies against the CH505 tier 2 TF virus.

6. Prime as in steps #1-4 above and then boost with the forms of the MT145 SIV Env (see e.g. Andrabi et al., 2019, Cell Reports 27, 2426–244) or similar SIV envelope that has a V1V2 loop –glycan bnAb epitope that binds to V1V2-glycan UCAs and bnAbs.
7. Prime as in steps #1-4 above and then boost with CM244, ZM233, WITO HIV-1 envelope or other WT Envs that have binding affinity for V1V2 bnAbs and their UCAs.

[0179] In non-limiting embodiments, these are administered as recombinant protein. Any suitable adjuvant could be use. In non-limiting embodiments, these are administered as nucleic acids, DNA and/or mRNAs. In non-limiting embodiments, the mRNAs are modified mRNAs administered as LNPs.

[0180] In non-limiting embodiments, the immunogens provide optimal prime for V1V2, and/or CD4 binding site, and/or Fusion Peptide precursors. In some embodiments, an optimal prime is determined by measurement of the frequency of bnAb precursors before immunization and after each immunization to determine if the immunization has expanded the desired bnAb B cell precursor pool. This can be performed by initial B cell repertoire analysis by single cell sorting of memory or germinal center B cells (e.g. Bonsignori et al. Sci Transl Med. 2017 Mar 15; 9(381): eaai7514.) and then followed by next generation sequencing of either lymph node, blood or other immune organ B cells to determine if the primed B cell bnAb clones were expanded and therefore boosted.

What is claimed is:

1. A recombinant HIV-1 envelope polypeptide from Table 1, Figures 3A-3C or Figures 4A-4B, wherein the polypeptide is a non-naturally occurring protomer designed to form an envelope trimer.
2. A nucleic acid encoding the recombinant polypeptide of claim 1.
3. A recombinant trimer comprising three identical protomers of an envelope from Table 1 Figures 3A-3C or Figures 4A-4B.
4. An immunogenic composition comprising the recombinant trimer of claim 3 and a carrier, wherein the trimer comprises three identical protomers of an HIV-1 envelope listed in Table 1, Figures 3A-3C or Figures 4A-4B.
5. An immunogenic composition comprising nucleic acid encoding the recombinant HIV-1 envelope of claim 1 and a carrier.
6. The composition of claim 4 or 5, further comprising an adjuvant.
7. The composition of any one of claim 2 or 5, wherein the nucleic acid is operably linked to a promoter, and wherein in certain embodiment the nucleic acid is inserted in an expression vector.
8. A method of inducing an immune response in a subject comprising administering a composition comprising any suitable form of a nucleic acid(s) of claim 2 or the polypeptide of claim 1 in an amount sufficient to induce an immune response.
9. The method of claim 8 wherein the nucleic acid encodes a gp120 envelope, gp120D8 envelope, a gp140 envelope (gp140C, gp140CF, gp140CFD) as soluble or stabilized protomer of a SOSIP trimer, a gp145 envelope, a gp150 envelope, a transmembrane bound envelope, or an envelope designed to multimerize.
10. The method of claim 8, wherein the polypeptide is gp120 envelope, gp120D8 envelope, a gp140 envelope (gp140C, gp140CF, gp140CFD) as soluble or stabilized protomer of a SOSIP trimer, a gp145 envelope, a gp150 envelope, a transmembrane bound envelope, or an envelope designed to multimerize.
11. The method of claim 8, wherein the composition further comprises an adjuvant.
12. The method of claim 8, further comprising administering an agent which modulates host immune tolerance.

13. The method of claim 10, wherein the polypeptide administered is multimerized in a liposome or nanoparticle.
14. The method of claim 8, further comprising administering one or more additional HIV-1 immunogens to induce a T cell response.
15. A composition comprising a nanoparticle and a carrier, wherein the nanoparticle comprises any one of the envelopes of claim 1.
16. The composition of claim 15, wherein the nanoparticle is ferritin self-assembling nanoparticle.
17. A composition comprising a nanoparticle and a carrier, wherein the nanoparticle comprises any one of the trimers of claims 2 or 3.
18. The composition of claim 17 wherein the nanoparticle is ferritin self-assembling nanoparticle.
19. The composition of claim 17, wherein the nanoparticle comprises multimers of trimers.
20. The composition of claim 17, wherein the nanoparticle comprises 1-8 trimers.
21. A method of inducing an immune response in a subject comprising administering an immunogenic composition comprising any one of the recombinant envelopes the preceding claims or compositions of the preceding claims.
22. The method of claim 21, wherein the composition is administered as a prime.
23. The method of claim 21, wherein the composition is administered as a boost.
24. A nucleic acid encoding any of the recombinant envelopes of the preceding claims.
25. A composition comprising the nucleic acid of claim 24 and a carrier.
26. A method of inducing an immune response in a subject comprising administering an immunogenic composition comprising the nucleic acid of claim 24 or the composition of claim 25.

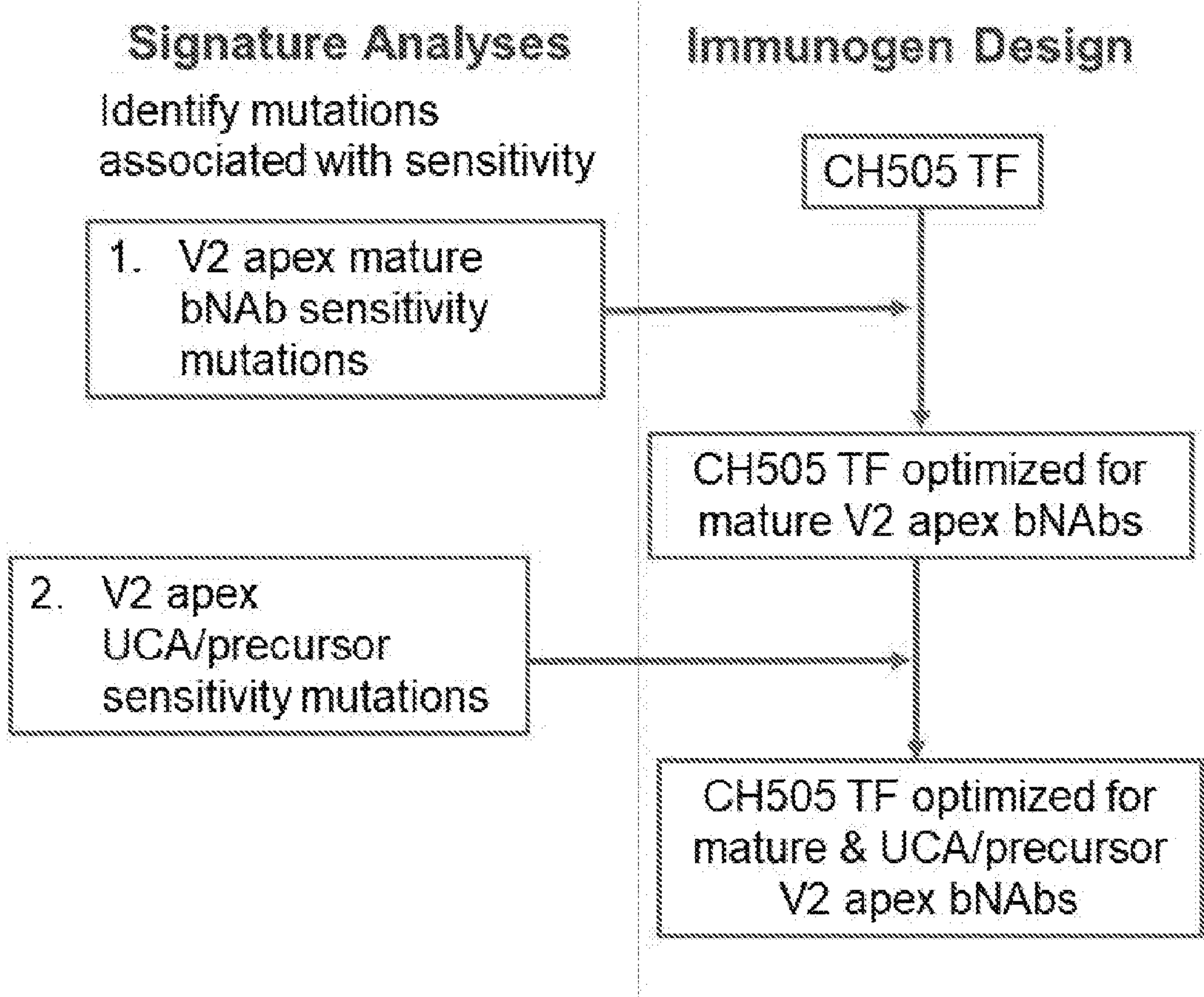


Figure 1A

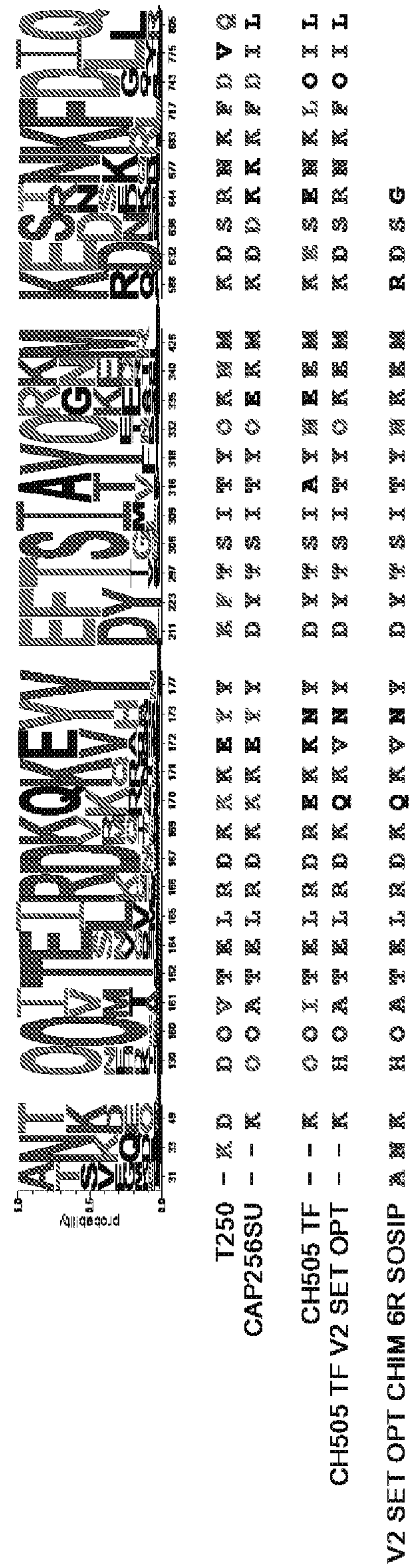


Figure 1B

Natural proteins for reference.

>CH0505.TF baseline natural immunogen

MRV-M---GIQRNYPQWWIWSMLGFWMLMI-CNG---MWVTVYYGVPVWKEAKTTLFCAS
DAKAYEKEVHNVWATHACVPTDPNPQEMVLKNVTENFNMWKNDMVDQMHEDEVISLWDQSL
KPCVKLTPLCVTLNCTNATAS-----NSSIIEGMKNCSFNITTELDRDKREKKNALF
YKLDIVQLDGN-----SSQYRLINCNTSVITQACPKVSFDEPIPIHYCAPAGYAILKC
NNKFTFTGTGPCNNVSTVQCTHGIKPVVSTQLLLNGLSLAEGEIIRSENI TNNVKTIIVHL
NESVKIECTRFNNKTRTSIRI--GPGQAFYATGQVIGDIREAYCNINESKWNETLQRVSK
KLKEYFP-HKNITFQPSSGGDLEITTHSFNCGGEFFYCNTSSLFNRTYMANSTDMA---N
STETNSTRTITIHCRIKQIINMWQEVGRAMYAPPIAGNITCISNITGLLLTRDGGKNN--
-TETFRPGGGNMKDNWRS ELYKYKVVEVKPLGVAPTNAARRRVVEREKR--AVGMGAVFLG
FLGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLKAIEAQQHMLKLTWVGIKQLQARVL
ALERYLKDQQLLGMWGC SGKLICTTNVYWNSSWSNKT YGDIWDNMTWMQWEREISNYTEI
IYELLEESQNQQEKNEQDLLALDRWNSLWNWFNITNWLWYIKIFIMIVGGLIGLRIIFAV
LSLVNRVRQGYSPLSLQTLIPSPRGPD RPGGIEEGGEQDRNRSTRLVSGFLALVWDDL R
SLCLFIYHRLRDFILIAARAGELLGRSSLKGLRRGWEALKYLGSLVQYWGLELKRS AISL
LDTLAI AVGEGTDRI LEFVLGICRAIRNIPTRIRQGFETALL*

>HXB2 reference strain

MRVKE---KYQHLWRWGWGWGTMLLG-MLMICSATEKLWVTVYYGVPVWKEATTTLFCAS
DAKAYDTEVHNVWATHACVPTDPNPQEVVLNVVTENFNMWKNDMVEQMHEDIISLWDQSL
KPCVKLTPLCVSLKCTDLKNDTNTNSSSGRMIMEKGEIKNCSFNISTIRGKVQKEYAFF
YKLDIIPIDN-----DTSYKLTSCNTSVITQACPKVSFDEPIPIHYCAPAGFAILKC
NNKTFNGTGPCNTVSTVQCTHGI RPVVSTQLLLNGLSLAE EEVVIRSVNFTDNAKTIIVQL
NTSVEINCTRPNNNTRKRI RIRQRGPGRAFVTIGK--IGNMRQAHCNISRAKWNNTLKQIAS
KLREQFGNNTIIFKQSSGGDPEIVTHSFNCGGEFFYCNTS QLENSTWFNSTW---STEG
SNNTEGSDTITLPCRIKQIINMWQKV GKAMYAPPI SGQIRCSSNITGLLLTRDGGNSN-N
ESEIFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPT KAKRRVVQREKR--AVGIGALFLG
FLGAAGSTMGAASMTLTVQARQLLSGIVQQQNLLRAIEAQQHLLQLTVWGIKQLQARIL
AVERYLKDQQLLGIWGC SGKLICTTAVPWNASWSNKSLEQIWNHTTWMEWDREINNYTSL
IHSLIEESQNQQEKNEQELLELDK WASLWNWFNITNWLWYIKL FIMIVGGLVGLRIVFAV
LSIVNRVRQGYSPLSFQTHLPTPRGPD RP EGIEEGGERDRDRSIRLVNGSLALI WDDL R
SLCLFSYHRLRDL LLLIVTRIVELLGR-----RGWEALKYWWNLLQYWSQELKNSAVSL
LNATAIAVAEGTDRVIEVVQGACRAIRHIPRIRQGLERILL*

>T250-4 natural V2 apex sensitive protein, V1V2 contributions

MRVMGIQRNYPPL---WRWGTMI FW-MMLCSA-EKLWVTVYYGVPVWREADTTLFCAS
DAKGYDTEAHNVWATHACVPTDPRPQEMYLENVTENFNMWKNSMVEQMHTDIISLWDESL
KPCVKLTPLCVTLDCQAFNSSSH-----TNSSIAMQEMKNCSFNVTTELDRKKKKEYSFF
YKTDIEQINK-----NGRQYRLINCNTSAITQACPKVSFDEPIPIHFCA PAGFAILKC
NEKHFNGKGPCKNVSTVQCTHGIKPVVSTQLLLNGLSLAE EEVVIRVENTIDNAKTIIVQL
AKPVKINCTRPNNNTRKSI RIR--GPGQTFYATGDIIGNIRKAYCNVSKREWNNTLQQVAA
QLSKSFNNTK-IVFEKHSGGDLEVI THSFVCGGEFFYCNTS GLENSTWTNSTWTN-STTG
SNGTESNDTITLQCEIKQFINMWQ RVGRAMYAPPI PGVIRCESDITGLLLTRDGPNST--
QNETFRPGGGDMRDNWRSELYKYKV VQIEPLGVAPTHAKRRRVVEREKR--AVGLGAVFFG
FLGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLKAIEAQQQLRLTVWGIKQLQARVL
ALERYLKDQQLLGIWGC SGKLICTTVPWNSSWSNKNYTDIWDNMTWLQWDREISNYTDE
IYRLIEQSQNQQEKNEQDLLALDKWASLWNWFDITNWLWYIKIFIMIVGGLIGLRIIFTV
LNVINRVRQGYSPLSFQTLTHHQREPDRPERIEEGGGEQDRDRSVRLVSGFLALAWDDL R
SLCLFSFHRLRDLVLIAARGVELLGHSS LKGLRLGWEALKLLGNLLSYWGQELKNSAINL
LDAVAIAVANWTD RVIKIGQRAGRAILNIPIRIRQGLERALL*

>ZM233.6 natural V2 apex sensitive protein, V1V2 contributions

MRVRGIMRNWQQW---WIWGS LGFW-MLIICNVMSLWVTVYYGVPVWREAKTTLFCAS
DAKAYETEHSVWATHACVPTDPNPQEMVLENVTENFNMWKNDMVDQMHEDEVIS IWDQSL
KPCVKLTPLCVTLDCSTYNN T-----HNISKEMKICSFNMTTELDRKKRKVNVL F
YKLDLVPLTNSS-----NTTNRYRLIS CNTSTITQACPKVSFDEPIPIHYCAPAGYAILKC

Figure 3A

NNKTFNGTGPCNNVSTVQCTHGIKPVVSTQLLNGSLAEEEEIIIRFENLTDNVKIIIVQL
NETINITCTRPNNNTRKSIIRI--GPGQSFYATGEIVGNI REAHCNISASKWNKTLERVRT
KLKEHFPNKT-IEFEPSSGGDLEITTHSFNCGGEFFYCNTSGLFNSAINGT-----
-----LTSNVTLPCRIKQIINMWQEVGRAMYAPPIAGNITCKSNITGLLLTRDGGENSSS
TTETFRPTGGDMKNNWRSELYKYKVVEIKPLGIAPTEAKRRVVEREKR--AVGIGAVFLG
FLGAAGSTMGAASMTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLQLTVWGIKQLQARVL
AIERYLKDQQLLGLWGCSGKLICTTNVPWNASWSNKSNDIWDNMTWMQWDREISNHTDT
IYRLLED SQNQQEKNEKDLLALDSWKNLWNWFSITKWLWYIKIFIMIVGGLIGLRIIFAV
LSIVNRVRQGYSPLSFQTLTPNPRGPDRLGGIEEEGGEQDKNKSRRLVGTGFLPVVWDDLRL
SLCLFSYHLLRDFILIVARTVELLGR-----RGWEALKYLGGLVQYWGLELKKSTISL
LDTIAIVVAEGTDRIIEVLQRIGRAIYNIPRIRQGFETALL*

>BG505.T332N used for gp41 stable SOSIP expression
MRVMGIQRNCQHL-----FRWGTMI LG-MIIICSAENLWVTVYYGVPVWKDAETTLFCAS
DAKAYETEKHNWATHACVPTDPNPQEIHLNVTEEFNMWKNMVEQMHTDIISLWDQSL
KPCVKLTPLCVTLQCTNVNITDDM-----RGELKNCSFNMTELRDKKQKVYSLF
YRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKC
KDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLNGSLAEEEVMI RSENI TNNAKNILVQF
NTPVQINCTRPNNNTRKSIIRI--GPGQAFYATGDIIGDIRQAHCNVSKATWNETLGKVVK
QLRKHFGNNTIIRFANSSGGDLEVTHSFNCGGEFFYCNTSGLFNSTWISNT----SVQG
SNSTGSNDSITLPCRIKQIINMWQRIQAMYPPIQGVI RCVSNITGLILTRDGGSTN-S
TTETFRPGGDMRDNRSELYKYKVVKIEPLGVAPTRAKRRVVGREKR--AVGIGAVFLG
FLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAIEAQQHLLKLTWGIKQLQARVL
AVERYLRDQQLLGIWGCSGKLICTTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQI
IYGLLEESQNQQEKNEQDLLALDKWASLWNWFDISNWLWYIKIFIMIVGGLIGLRIIFAV
LSVIHRVRQGYSPLSFQTHTPNPRGLDRPERIEEEDGEQDRGRSTRLVSGFLALAWDDLRL
SLCLFCYHRLRDFILIAARIVELLGHSSLKGLRLGWGLKYLWNLLAYWGRELKISAINL
EDTIAIAVAEWTDRIIEIGQLCRAFLHIPRIRQGLERALL*

Figure 3A continued

CH505 Full length modified proteins.

>V2.SET.OPT
MRV-M---GIQRNYPQWWIWSMLGFWMLMI--CNG---MWVTVYYGVPVWKEAKTTLFCAS
DAKAYEKEVHNVWATHACVPTDPNPQEMVLKNVTENFNMWKNDMVDQMHEDEVISLWDQSL
KPCVKLTPLCVTLHCSTYNNT-----HNISKGMKNCSFNATTEL RDKKQKVNALF
YKLDIVQLNK-----NGRQYRLINCNTSVITQACPKVSFDPIPIHYCAPAGYAILKC
NNKTFNGTGPCNNVSTVQCTHGIKPVVSTQLLLNGSLAEGEIIIRSENI TNNVKTIIVHL
NESVKIECTRPNNKTRTSIRI--GPGQTFYATGQVIGDIREAYCNINKSKWNETLQRVSK
KLKEYFP-HKNITFQPSSGGDLEITTHSFNCGGEFFYCNTSSLFNRTYMANSTDMA---N
STETNSTRTITIHCRIKQIINMWQEVGRAMYAPPIAGNITCISNITGLLLTRDGGKNN---
-TETFRPGGGNMKDNWRS ELYKYKVVEVKPLGVAPT NARRRVVEREKR--AVGMGAVFLG
FLGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLKAIEAQQHMLKLTWGIKQLQARVL
ALERYLKDQQLLGMWGCSGKLICTTNVYWNSSWSNKTYGDIWDNMTWMQWDREISNYTEI
IYRLLEESQNQQEKNEQDLLALDRWNSLWNWFNITNWLWYIKIFIMIVGGLIGLRIIFAV
LSLVNVRVQGYSPLSFQTLIPSPRGPDRPGGIEEGGEQDNRSTRLVSGFLALVWDDLRL
SLCLFIYHRLRDFILIAARAGELLGRSSLKGLRRGWEALKYLGSLVQYWGLELKRS AISL
LDTLAIAVGEGTDRILEFVLGICRAIRNIPTRIRQGFETALL*

>V2.SET.OPT.N332
MRV-M---GIQRNYPQWWIWSMLGFWMLMI--CNG---MWVTVYYGVPVWKEAKTTLFCAS
DAKAYEKEVHNVWATHACVPTDPNPQEMVLKNVTENFNMWKNDMVDQMHEDEVISLWDQSL
KPCVKLTPLCVTLHCSTYNNT-----HNISKGMKNCSFNATTEL RDKKQKVNALF
YKLDIVQLNK-----NGRQYRLINCNTSVITQACPKVSFDPIPIHYCAPAGYAILKC
NNKTFNGTGPCNNVSTVQCTHGIKPVVSTQLLLNGSLAEGEIIIRSENI TNNVKTIIVHL
NESVKIECTRPNNKTRTSIRI--GPGQTFYATGQVIGDIREAYCNINKSKWNETLQRVSK
KLKEYFP-HKNITFQPSSGGDLEITTHSFNCGGEFFYCNTSSLFNRTYMANSTDMA---N
STETNSTRTITIHCRIKQIINMWQEVGRAMYAPPIAGNITCISNITGLLLTRDGGKNN---
-TETFRPGGGNMKDNWRS ELYKYKVVEVKPLGVAPT NARRRVVEREKR--AVGMGAVFLG
FLGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLKAIEAQQHMLKLTWGIKQLQARVL
ALERYLKDQQLLGMWGCSGKLICTTNVYWNSSWSNKTYGDIWDNMTWMQWDREISNYTEI
IYRLLEESQNQQEKNEQDLLALDRWNSLWNWFNITNWLWYIKIFIMIVGGLIGLRIIFAV
LSLVNVRVQGYSPLSFQTLIPSPRGPDRPGGIEEGGEQDNRSTRLVSGFLALVWDDLRL
SLCLFIYHRLRDFILIAARAGELLGRSSLKGLRRGWEALKYLGSLVQYWGLELKRS AISL
LDTLAIAVGEGTDRILEFVLGICRAIRNIPTRIRQGFETALL*

>V2.UCA.OPT1
MRV-M---GIQRNYPQWWIWSMLGFWMLMI--CNG---MWVTVYYGVPVWKEAKTTLFCAS
DAKAYEKEVHNVWATHACVPTDPNPQEMVLKNVTENFNMWKNDMVDQMHEDEVISLWDQSL
KPCVKLTPLCVTLHCSTYNNT-----HNISKGMKNCSFNMTTEL RDKKQKVNALF
YKLDIVQLNK-----NGRQYRLINCNTSAITQACPKVSFDPIPIHYCAPAGYAILKC
NNKTFNGTGPCNNVSTVQCTHGIKPVVSTQLLLNGSLAEGEIIIRSENI TNNVKTIIVHL
NESVKIECTRPNNKTRKSIRI--GPGQTFYATGDVIGDIREAYCNINKSKWNETLQRVSK
KLKEYFP-HKNITFQPSSGGDLEITTHSFNCGGEFFYCNTSSLFNRTYMANSTDMA---N
STETNSTRTITIHCRIKQIINMWQEVGRAMYAPPIAGNITCISNITGLLLTRDGGKNN---
-TETFRPGGGNMKDNWRS ELYKYKVVEVKPLGVAPT NARRRVVEREKR--AVGMGAVFLG
FLGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLKAIEAQQHMLKLTWGIKQLQARVL
ALERYLKDQQLLGMWGCSGKLICTTNVYWNSSWSNKTYGDIWDNMTWMQWDREISNYTEI
IYRLLEESQNQQEKNEQDLLALDRWNSLWNWFNITNWLWYIKIFIMIVGGLIGLRIIFAV
LSLVNVRVQGYSPLSFQTLIPSPRGPDRPGRIEEGGEQDNRSTRLVSGFLALVWDDLRL
SLCLFIYHRLRDFILIAARAGELLGRSSLKGLRRGWEALKYLGSLVQYWGLELKRS AISL
LDTLAIAVGEGTDRILEFVLGICRAIRNIPTRIRQGFETALL*

>V2.UCA.OPT1.N332
MRV-M---GIQRNYPQWWIWSMLGFWMLMI--CNG---MWVTVYYGVPVWKEAKTTLFCAS
DAKAYEKEVHNVWATHACVPTDPNPQEMVLKNVTENFNMWKNDMVDQMHEDEVISLWDQSL
KPCVKLTPLCVTLHCSTYNNT-----HNISKGMKNCSFNMTTEL RDKKQKVNALF

Figure 3B

YKLDIVQLNK-----NGRQYRLINCNTSAITQACPKVSFDPPIPIHYCAPAGYAILKC
 NNKTFNGTGPCNNVSTVQCTHGIKPVVSTQLLNGSLAEGEIIRSENI TNNVKTIIVHL
 NESVKIECTRPNNKTRKSI RI--GPGQTFYATGDVIGDI REAYCNISKS KWN ET LQ RVSK
 KLKEYFP-HKNITFQPSSGGDLEITTHSFNCGGEFFYCNTSSLENRTYMANSTDMA---N
 STETNSTRITITIHCRIKQI INMWQEVGRAMYAPPIAGNITCISNITGLLLTRDGGKNN--
 -TETFRPGGGNMKDNWRSELYKYKVVEVKPLGVAPTNAARRRVVEREKR--AVGMGAVFLG
 FLGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLKAIEAQQHMLKLTWVGIKQLQARVL
 ALERYLKDQQLLGMWGCSGKLICTTNVYWNSSWSNKT YGDIWDNM TWMQWDREI SNYTEI
 IYRLLEESQNQQEKNEQD L LALDRWNSLWNWFNITNWLWYIKIFIMIVGGLIGLRIIFAV
 LSLVNRVRQGYSPLSFQTLIPSPRGPDRPGRIEEGGGEQDRNRSTRLVSGFLALVWDDL R
 SLCLFIYHRLRDFILIAARAGELLGRSSLKGLRRGWEALKYLGSLVQYWGLELKRSAISL
 LDTLAI AVGEGTDRILEFVLGICRAIRNIPTRI RQGFETALL*

>V2.UCA.OPT2

MRV-M---GIQRNYPQWWIWSMLGFWMLMI--CNG---MWVTVYYGVPVWKEAKTTLFCAS
 DAKAYEKEVHNVWATHACVPTDPNPQEMVLKNVTENFNMWKNDMVDQMHEDEVISLWDQSL
 KPCVKLTPLCVTLHCSTYNN-----HNISKGMKNCSFNMTTEL RDKKQKVNALF
 YKLDIVQLNK-----NGRQYRLINCNTSAITQACPKVSFDPPIPIHYCAPAGYAILKC
 NNKTFNGTGPCNNVSTVQCTHGIKPVVSTQLLNGSLAEGEIIRSENI TNNVKTIIVHL
 NESVKIECTRPNNKTRKSI RI--GPGQTFYATGDVIGDI REAYCNINKS KWN ET LQ RVSK
 KLKEYFP-HKNITFQPSSGGDLEITTHSFNCGGEFFYCNTSSLENRTYMANSTDMA---N
 STETNSTRITITIHCRIKQI INMWQEVGRAMYAPPIAGNITCISNITGLLLTRDGGKNN--
 -TETFRPGGGNMKDNWRSELYKYKVVEVKPLGVAPTNAARRRVVEREKR--AVGMGAVFLG
 FLGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLKAIEAQQHMLKLTWVGIKQLQARVL
 ALERYLKDQQLLGMWGCSGKLICTTNVYWNSSWSNKT YGDIWDNM TWMQWDREI SNYTEI
 IYRLLEESQNQQEKNEQD L LALDRWNSLWNWFNITNWLWYIKIFIMIVGGLIGLRIIFAV
 LSLVNRVRQGYSPLSFQTLIPSPRGPDRPGRIEEGGGEQDRNRSTRLVSGFLALVWDDL R
 SLCLFIYHRLRDFILIAARAGELLGRSSLKGLRRGWEALKYLGSLVQYWGLELKRSAISL
 LDTLAI AVGEGTDRILEFVLGICRAIRHIPTRIRQGFETALL*

>V2.UCA.OPT2.N332

MRV-M---GIQRNYPQWWIWSMLGFWMLMI--CNG---MWVTVYYGVPVWKEAKTTLFCAS
 DAKAYEKEVHNVWATHACVPTDPNPQEMVLKNVTENFNMWKNDMVDQMHEDEVISLWDQSL
 KPCVKLTPLCVTLHCSTYNN-----HNISKGMKNCSFNMTTEL RDKKQKVNALF
 YKLDIVQLNK-----NGRQYRLINCNTSAITQACPKVSFDPPIPIHYCAPAGYAILKC
 NNKTFNGTGPCNNVSTVQCTHGIKPVVSTQLLNGSLAEGEIIRSENI TNNVKTIIVHL
 NESVKIECTRPNNKTRKSI RI--GPGQTFYATGDVIGDI REAYCNISKS KWN ET LQ RVSK
 KLKEYFP-HKNITFQPSSGGDLEITTHSFNCGGEFFYCNTSSLENRTYMANSTDMA---N
 STETNSTRITITIHCRIKQI INMWQEVGRAMYAPPIAGNITCISNITGLLLTRDGGKNN--
 -TETFRPGGGNMKDNWRSELYKYKVVEVKPLGVAPTNAARRRVVEREKR--AVGMGAVFLG
 FLGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLKAIEAQQHMLKLTWVGIKQLQARVL
 ALERYLKDQQLLGMWGCSGKLICTTNVYWNSSWSNKT YGDIWDNM TWMQWDREI SNYTEI
 IYRLLEESQNQQEKNEQD L LALDRWNSLWNWFNITNWLWYIKIFIMIVGGLIGLRIIFAV
 LSLVNRVRQGYSPLSFQTLIPSPRGPDRPGRIEEGGGEQDRNRSTRLVSGFLALVWDDL R
 SLCLFIYHRLRDFILIAARAGELLGRSSLKGLRRGWEALKYLGSLVQYWGLELKRSAISL
 LDTLAI AVGEGTDRILEFVLGICRAIRHIPTRIRQGFETALL*

>CH505.TF.chim.6R.SOSIP

MPM-----GSLQPLATLYLL---GMLVAS--VLA AENLWVTVYYGVPVWKEAKTTLFCAS
 DAKAYEKKVHNVWATHACVPTDPNPQEMVLKNVTENFNMWKNDMVDQMHEDEVISLWDQSL
 KPCVKLTPLCVTLNCTNATAS-----NSSIIEGMKNCSFNITTEL RDKREKKNALF
 YKLDIVQLDGN-----SSQYRLINCNTSVITQACPKVSFDPPIPIHYCAPAGYAILKC
 NNKTFNGTGPCNNVSTVQCTHGIKPVVSTQLLNGSLAEGEIIRSENI TNNVKTIIVHL
 NESVKIECTRPNNKTRTSI RI--GPGQWFYATGQVIGDI REAYCNINES KWN ET LQ RVSK
 KLKEYFP-HKNITFQPSSGGDLEITTHSFNCGGEFFYCNTSSLENRTYMANSTDMA---N
 STETNSTRITITIHCRIKQI INMWQEVGRAMYAPPIAGNITCISNITGLLLTRDGGKNN--
 -TETFRPGGGNMKDNWRSELYKYKVVKIEPLGVAPT RCKRRRVVGRRRRRRAVGI GAVFLG
 FLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQQHLLKLTWVGIKQLQARVL
 AVERYLRDQQLLGIWGCSGKLICTTNVPWNSSWSNPNLSEIWDNM T W LQWDKEI SNYTQI
 IYGLLEESQNQQEKNEQD L LALD*

Figure 3B continued

Modified proteins revised for expression as SOSiP proteins.

>CH505TF_V2.SET.OPT_ch.SOSiPv4.1 (A)
MPM-----GSLQPLATLYLL---GMLVAS---VLAAENLWVTVYYGVPVWKEAKTTLFCAS
DAKAYEKKVHNVWATHACVPTDPNPQEMVLKNVTENFNMWKNMVDQMHEDEVISLWDQSL
KPCVKLTPLCVTLHCSTYNNT-----HNISKGMKNCSFNATTEL RDKKQKVNALF
YKLDIVQLNK-----NGRQYRLINCNTSVITQACPKVSFDPIPIHYCAPAGYAILKC
NNKTFNGTGPCNNVSTVQCTHGIKPVVSTQLLLNGSLAEGEIIIRSENI TNNVKTIIVHL
NESVKIECTRPNNKTRTSIRI--GPGQTFYATGQVIGDIREAYCNINKSKWNETLQRVSK
KLKEYFP-HKNITFQPSSGGDLEITTHSFNCGGEFFYCNTSSLFNRTYMANSTDMA---N
STETNSTRITITIHCRIKQIINMWQEVGRAMYAPPIAGNITCISNITGLLLTRDGGKNN---
-TETFRPGGGMKDNWRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVFLG
FLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQQHLLKLTWVGIKQLQARVL
AVERYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNPNLSEIWDNMTWLQWDKEISNYTQI
IYGLLEESQNQQEKNEQDILLALD*

>CH505TF_V2.SET.OPT.N332_ch.SOSiPv4.1 (B)
MPM-----GSLQPLATLYLL---GMLVAS---VLAAENLWVTVYYGVPVWKEAKTTLFCAS
DAKAYEKKVHNVWATHACVPTDPNPQEMVLKNVTENFNMWKNMVDQMHEDEVISLWDQSL
KPCVKLTPLCVTLHCSTYNNT-----HNISKGMKNCSFNATTEL RDKKQKVNALF
YKLDIVQLNK-----NGRQYRLINCNTSVITQACPKVSFDPIPIHYCAPAGYAILKC
NNKTFNGTGPCNNVSTVQCTHGIKPVVSTQLLLNGSLAEGEIIIRSENI TNNVKTIIVHL
NESVKIECTRPNNKTRTSIRI--GPGQTFYATGQVIGDIREAYCNINKSKWNETLQRVSK
KLKEYFP-HKNITFQPSSGGDLEITTHSFNCGGEFFYCNTSSLFNRTYMANSTDMA---N
STETNSTRITITIHCRIKQIINMWQEVGRAMYAPPIAGNITCISNITGLLLTRDGGKNN---
-TETFRPGGGMKDNWRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVFLG
FLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQQHLLKLTWVGIKQLQARVL
AVERYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNPNLSEIWDNMTWLQWDKEISNYTQI
IYGLLEESQNQQEKNEQDILLALD*

>CH505TF_V2.UCA.OPT1_ch.SOSiPv4.1 (C)
MPM-----GSLQPLATLYLL---GMLVAS---VLAAENLWVTVYYGVPVWKEAKTTLFCAS
DAKAYEKKVHNVWATHACVPTDPNPQEMVLKNVTENFNMWKNMVDQMHEDEVISLWDQSL
KPCVKLTPLCVTLHCSTYNNT-----HNISKGMKNCSFNMTTEL RDKKQKVNALF
YKLDIVQLNK-----NGRQYRLINCNTSAITQACPKVSFDPIPIHYCAPAGYAILKC
NNKTFNGTGPCNNVSTVQCTHGIKPVVSTQLLLNGSLAEGEIIIRSENI TNNVKTIIVHL
NESVKIECTRPNNKTRKSIRI--GPGQTFYATGDVIGDIREAYCNINKSKWNETLQRVSK
KLKEYFP-HKNITFQPSSGGDLEITTHSFNCGGEFFYCNTSSLFNRTYMANSTDMA---N
STETNSTRITITIHCRIKQIINMWQEVGRAMYAPPIAGNITCISNITGLLLTRDGGKNN---
-TETFRPGGGMKDNWRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVFLG
FLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQQHLLKLTWVGIKQLQARVL
AVERYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNPNLSEIWDNMTWLQWDKEISNYTQI
IYGLLEESQNQQEKNEQDILLALD*

>CH505TF_V2.UCA.OPT1.N332_ch.SOSiP.v4.1 (D)
MPM-----GSLQPLATLYLL---GMLVAS---VLAAENLWVTVYYGVPVWKEAKTTLFCAS
DAKAYEKKVHNVWATHACVPTDPNPQEMVLKNVTENFNMWKNMVDQMHEDEVISLWDQSL
KPCVKLTPLCVTLHCSTYNNT-----HNISKGMKNCSFNMTTEL RDKKQKVNALF
YKLDIVQLNK-----NGRQYRLINCNTSAITQACPKVSFDPIPIHYCAPAGYAILKC
NNKTFNGTGPCNNVSTVQCTHGIKPVVSTQLLLNGSLAEGEIIIRSENI TNNVKTIIVHL
NESVKIECTRPNNKTRKSIRI--GPGQTFYATGDVIGDIREAYCNINKSKWNETLQRVSK
KLKEYFP-HKNITFQPSSGGDLEITTHSFNCGGEFFYCNTSSLFNRTYMANSTDMA---N
STETNSTRITITIHCRIKQIINMWQEVGRAMYAPPIAGNITCISNITGLLLTRDGGKNN---
-TETFRPGGGMKDNWRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVFLG
FLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQQHLLKLTWVGIKQLQARVL
AVERYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNPNLSEIWDNMTWLQWDKEISNYTQI
IYGLLEESQNQQEKNEQDILLALD*

Figure 3C

>CH505TF_V2.UCA.OPT2_ch.SOSIPv4.1 (E)
MPM-----GSLQPLATLYLL---GMLVAS--VLAAENLWVTVYYGVPVWKEAKTTLFCAS
DAKAYEKKVHNVWATHACVPTDPNPQEMVLKNVTENFNMWKNDMVDQMHEDEVISLWDQSL
KPCVKLTPLCVTLHCSTYNNT-----HNISKGMKNCSFNMTTEL RDKKQKVNALF
YKLDIVQLNK-----NGRQYRLINCNTSAITQACPKVSFDPIPIHYCAPAGYAILKC
NNKTFNGTGPCNNVSTVQCTHGIKPVVSTQLLNGSLAEGEIIIRSENI TNNVKTIIIVHL
NESVKIECTRPNNKTRKSIRI---GPGQTFYATGDVIGDI REAYCNINKSKWNETLQRVSK
KLKEYFP-HKNITFQPSGGDLEITTHSFNCGGEFFYCNTSSLENRTYMANSTDMA---N
STETNSTRITITIHCRIKQIINMWQEVGRAMYAPPIAGNITCISNITGLLLTRDGGKNN--
-TETFRPGGGNMKDNWRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRAVGIGAVFLG
FLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQQHLLKLTWVGIKQLQARVL
AVERYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQI
IYGLLEESQNQQEKNEQDLLALD*

>CH505TF_V2.UCA.OPT2.N332_ch.SOSIP.v4.1 (F)
MPM-----GSLQPLATLYLL---GMLVAS--VLAAENLWVTVYYGVPVWKEAKTTLFCAS
DAKAYEKKVHNVWATHACVPTDPNPQEMVLKNVTENFNMWKNDMVDQMHEDEVISLWDQSL
KPCVKLTPLCVTLHCSTYNNT-----HNISKGMKNCSFNMTTEL RDKKQKVNALF
YKLDIVQLNK-----NGRQYRLINCNTSAITQACPKVSFDPIPIHYCAPAGYAILKC
NNKTFNGTGPCNNVSTVQCTHGIKPVVSTQLLNGSLAEGEIIIRSENI TNNVKTIIIVHL
NESVKIECTRPNNKTRKSIRI---GPGQTFYATGDVIGDI REAYCNINKSKWNETLQRVSK
KLKEYFP-HKNITFQPSGGDLEITTHSFNCGGEFFYCNTSSLENRTYMANSTDMA---N
STETNSTRITITIHCRIKQIINMWQEVGRAMYAPPIAGNITCISNITGLLLTRDGGKNN--
-TETFRPGGGNMKDNWRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRAVGIGAVFLG
FLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQQHLLKLTWVGIKQLQARVL
AVERYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQI
IYGLLEESQNQQEKNEQDLLALD*

>CH505TF_V2.UCA.OPT1_gp41mut_ch.SOSIP.v4.1 (G)
MPM-----GSLQPLATLYLL---GMLVAS--VLAAENLWVTVYYGVPVWKEAKTTLFCAS
DAKAYEKKVHNVWATHACVPTDPNPQEMVLKNVTENFNMWKNDMVDQMHEDEVISLWDQSL
KPCVKLTPLCVTLHCSTYNNT-----HNISKGMKNCSFNMTTEL RDKKQKVNALF
YKLDIVQLNK-----NGRQYRLINCNTSAITQACPKVSFDPIPIHYCAPAGYAILKC
NNKTFNGTGPCNNVSTVQCTHGIKPVVSTQLLNGSLAEGEIIIRSENI TNNVKTIIIVHL
NESVKIECTRPNNKTRKSIRI---GPGQTFYATGDVIGDI REAYCNINKSKWNETLQRVSK
KLKEYFP-HKNITFQPSGGDLEITTHSFNCGGEFFYCNTSSLENRTYMANSTDMA---N
STETNSTRITITIHCRIKQIINMWQEVGRAMYAPPIAGNITCISNITGLLLTRDGGKNN--
-TETFRPGGGNMKDNWRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRAVGIGAVFLG
FLGAAGSTMGAASITLTVQARNLLSGIVQQQSNLLRAPEAQQHLLKLTWVGIKQLQARVL
AVERYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQI
IYRLLEESQNQQEKNEQDLLALD*

>CH505TF_V2.UCA.OPT1.N332_gp41mut_ch.SOSIP.v4.1 (H)
MPM-----GSLQPLATLYLL---GMLVAS--VLAAENLWVTVYYGVPVWKEAKTTLFCAS
DAKAYEKKVHNVWATHACVPTDPNPQEMVLKNVTENFNMWKNDMVDQMHEDEVISLWDQSL
KPCVKLTPLCVTLHCSTYNNT-----HNISKGMKNCSFNMTTEL RDKKQKVNALF
YKLDIVQLNK-----NGRQYRLINCNTSAITQACPKVSFDPIPIHYCAPAGYAILKC
NNKTFNGTGPCNNVSTVQCTHGIKPVVSTQLLNGSLAEGEIIIRSENI TNNVKTIIIVHL
NESVKIECTRPNNKTRKSIRI---GPGQTFYATGDVIGDI REAYCNINKSKWNETLQRVSK
KLKEYFP-HKNITFQPSGGDLEITTHSFNCGGEFFYCNTSSLENRTYMANSTDMA---N
STETNSTRITITIHCRIKQIINMWQEVGRAMYAPPIAGNITCISNITGLLLTRDGGKNN--
-TETFRPGGGNMKDNWRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRAVGIGAVFLG
FLGAAGSTMGAASITLTVQARNLLSGIVQQQSNLLRAPEAQQHLLKLTWVGIKQLQARVL
AVERYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQI
IYRLLEESQNQQEKNEQDLLALD*

Figure 3C continued

Gene number	Protein name
HV1301908	CH505TF_V2.UCA.OPT1.gp41mut_ch.SOSIP.v4.1
HV1301909	CH505TF_V2.UCA.OPT1.N332.gp41mut_ch.SOSIP.v4.1
HV1301910	CH505TF_V2.SET.OPT_ch.SOSIPv4.1
HV1301911	CH505TF_V2.SET.OPT.N332_ch.SOSIPv4.1
HV1301912	CH505TF_V2.UCA.OPT1_ch.SOSIPv4.1
HV1301913	CH505TF_V2.UCA.OPT1.N332_ch.SOSIP.v4.1

Nucleic Acid sequences

>HV1301908 (CH505TF_V2.UCA.OPT1.gp41mut_ch.SOSIP.v4.1)

GTCGACGCCACCATGCCATATGGGATCTCTGCAGCCCTCTGGCCACACTGTACCTGCTGGGAATGCTGGTGGCTT
CTSTGCTGGCCGCGGAGAATCTGTGGGTACAGTGTACTATGGCGTGCCCGTGTGGAAAGAGGCCAAGACCA
CACTGTTCTGTGCCAGCGACGCCAAGGCCTACGAGAAGAAAGTGCACAACGTGTGGGCCACTCACGCCTGCG
TTCCACCGATCCTAATCCTCAAGAGATGGTGTGAAGAAGCTGACCGAGAAGTTCAACATGTGGAAGAAGC
ACATGGTGGACCAGATGCACGAGGACGTGATCAGCCTGTGGGACCAGAGCCTGAAGCCTTGCGTGAAGCTG
ACCCCTCTGTGTGTGACCCTGCACTGCAGCACCTACAACAACACCCACAACATCAGCAAGGGCATGAAGAACT
GCAGCTTCAATATGACCACCGAGCTGCGGGACAAGAAGCAGAAAGTCAACGCCCTGTTCTACAAGCTGGACA
TCGTGCAGCTGAACAAGAACGGCCGGCAGTACCGGCTGATCAACTGCAACACAAGCGCCATCACACAGGCTT
GCCCCAAGGTGTCCTTCGATCCCATTCCTATCCACTACTGTGCCCTGCCGGCTACGCCATCCTGAAGTGCAAC
AACAAGACCTTCAACGGCACAGGCCCTGCAACAACGTGTCCACCGTGAGTGTACCCACGGCATCAAGCCAG
TGGTGTCTACCCAGCTGCTGCTGAATGGATCTCTGGCCGAGGGCGAGATCATCATCAGAAGCGAGAACATCA
CCAACAACGTCAAGACCATCATCGTCCACCTGAACGAGAGCGTGAAGATCGAGTGCACCCGGCCTAACAACA
AAACCAGAAAGAGCATCCGGATCGGCCCTGGCCAGACCTTTTATGCTACTGGCGACGTGATCGGCGACATCA
GAGAGGCCCTACTGCAACATCAACAAGAGCAAGTGGAACGAGACACTGCAGCGGGTGTCCAAGAACTGAAA
GAGTACTTCCCGCACAGAATATCACCTTCCAGCCTAGCTCTGGCGGCGACCTGGAAATCACCACACACAGCT
TTAACTGTGGCGGCGAGTTCTTCTACTGCAATACCAGCAGCCTGTTCAACCGGACCTACATGGCCAACTCCACC
GATATGGCCAACAGCACCGAGACAAACAGCACCCGGACCATCACCATCCACTGCCGGATCAAGCAGATCATC
AATATGTGGCAAGAAGTCGGCAGGGCTATGTACGCCCTCCTATCGCCGGCAACATCACCTGTATCAGCAATA
TCACCGGCCTGCTGCTCACCAGAGATGGCGGCAAGAACAACACCGAAACCTTCAGACCCGGCGGAGGCAACA
TGAAGGACAATTGGAGAAGCGAGCTGTACAAGTACAAGGTGGTCAAGATCGAGCCCCTGGGCGTCGCACCA
ACACGGTGCAAGAGAAGAGTCGTGGGCGGTCGTAGAAGGCGGAGAGCCGTTGGAATTGGCGCCGTGTTCTT
GGGCTTTCTGGGAGCCGCTGGATCTACAATGGGCGCTGCCAGCATCACCTGACAGTGCAGGCTAGAAATCT
GCTGAGCGGCATTGTGCAGCAGCAGAGCAACCTGCTGAGAGCCCCTGAAGCTCAGCAGCACCT

Figure 4A

CCTGAAACTGACCGTGTGGGGAATCAAGCAGCTGCAGGCAAGAGTGCTGGCAGTGGAAAGATACCTGAAGG
ACCAGCAGCTCCTCGGAATCTGGGGCTGTTCTGGCAAGCTGATCTGCTGCACCAACGTGCCCTGGAACAGCTC
CTGGTCCAACAGAAACCTGAGCGAGATCTGGGACAACATGACCTGGCTGCAGTGGGACAAAGAGATCAGCA
ACTACACCCAGATCATCTACCGGCTGCTGGAAGAGAGCCAGAACCAGCAAGAGAAAAACGAGCAGGACCTGC
TGGCCCTGGACTGATAAGGATCC

>HV1301909 (CH505TF_V2.UCA.OPT1.N332.gp41mut_ch.SOSIP.v4.1)

GTCGACGCCACCATGCCTATGGGATCTCTGCAGCCTCTGGCCACACTGTACCTGCTGGGAATGCTGGTGGCTT
CTGTGCTGGCCGCGAGAATCTGTGGGTACAGTGTACTATGGCGTGCCCGTGTGGAAAGAGGCCAAGACCA
CACTGTTCTGTGCCAGCGACGCCAAGGCCTACGAGAAGAAAGTGCAACAGTGTGGGCCACTCACGCCTGCG
TTCCCACCGATCCTAATCCTCAAGAGATGGTGCTGAAGAACGTGACCGAGAACTTCAACATGTGGAAGAACG
ACATGGTGGACCAGATGCACGAGGACGTGATCAGCCTGTGGGACCAGAGCCTGAAGCCTTGCGTGAAGCTG
ACCCCTCTGTGTGTGACCCTGCACTGCAGCACCTACAACAACACCCACAACATCAGCAAGGGCATGAAGAACT
GCAGCTTCAATATGACCACCGAGCTGCGGGACAAGAAGCAGAAAAGTCAACGCCCTGTTCTACAAGCTGGACA
TCGTGCAGCTGAACAAGAACGGCCGGCAGTACCGGCTGATCAACTGCAACACAAGCGCCATCACACAGGCTT
GCCCCAAGGTGTCCTTCGATCCCATTCCTATCCACTACTGTGCCCTGCCGGCTACGCCATCCTGAAGTGCAAC
AACAAGACCTTCAACGGCACAGGCCCTGCAACAACGTGTCCACCGTGCAAGTGTACCCACGGCATCAAGCCAG
TGGTGTCTACCCAGCTGCTGCTGAATGGATCTCTGGCCGAGGGCGAGATCATCATCAGAAGCGAGAACATCA
CCAACAACGTCAAGACCATCATCGTCCACCTGAACGAGAGCGTGAAGATCGAGTGACCCCGGCCTAACAACA
AAACCAGAAAGAGCATCCGGATCGGCCCTGGCCAGACCTTTTATGCTACTGGCGACGTGATCGGCGACATCA
GAGAGGCCTACTGCAACATCTCAAGAGCAAGTGGAACGAGACACTGCAGCGGGTGTCCAAGAACTGAAA
GAGTACTTCCCGCACAGAATATCACCTTCCAGCCTAGCTCTGGCGGCGACCTGGAAATCACCACACACAGCT
TTAACTGTGGCGGCGAGTTCTTCTACTGCAATACCAGCAGCCTGTTCAACCGGACCTACATGGCCAACTCCACC
GATATGGCCAACAGCACCGAGACAAACAGCACCCGGACCATCACCATCCACTGCCGGATCAAGCAGATCATC
AATATGTGGCAAGAAGTCGGCAGGGCTATGTACGCCCTCCTATCGCCGGCAACATCACCTGTATCAGCAATA
TCACCGGCCTGCTGCTCACCAGAGATGGCGGCAAGAACAACACCGAAACCTTCAGACCCGGCGGAGGCAACA
TGAAGGACAATTGGAGAAGCGAGCTGTACAAGTACAAGGTGGTCAAGATCGAGCCCCTGGGCGTCGCACCA
ACACGGTGCAAGAGAAGAGTCGTGGGCCGTCGTAGAAGGCGGAGAGCCGTTGGAATTGGCGCCGTGTTCTT
GGGCTTTCTGGGAGCCGCTGGATCTACAATGGGCGCTGCCAGCATCACCTGACAGTGCAGGCTAGAAATCT
GCTGAGCGGCATTGTGCAGCAGCAGAGCAACCTGCTGAGAGCCCCTGAAGCTCAGCAGCACCTCCTGAACT
GACCGTGTGGGGAATCAAGCAGCTGCAGGCAAGAGTGCTGGCAGTGGAAAGATACCTGAAGGACCAGCAGC
TCCTCGGAATCTGGGGCTGTTCTGGCAAGCTGATCTGCTGCACCAACGTGCCCTGGAACAGCTCCTGGTCCAA
CAGAAACCTGAGCGAGATCTGGGACAACATGACCTGGCTGCAGTGGGACAAAGAGATCAGCAACTACACCCA
GATCATCTACCGGCTGCTGGAAGAGAGCCAGAACCAGCAAGAGAAAAACGAGCAGGACCTGCTGGCCCTGG
ACTGATAAGGATCC

>HV1301910 (CH505TF_V2.SET.OPT_ch.SOSIPv4.1)

GTCGACGCCACCATGCCTATGGGATCTCTGCAGCCTCTGGCCACACTGTACCTGCTGGGAATGCTGGTGGCTT
CTGTGCTGGCCGCGAGAATCTGTGGGTACAGTGTACTATGGCGTGCCCGTGTGGAAAGAGGCCAAGACCA
CACTGT

Figure 4A continued

TCTGTGCCAGCGACGCCAAGGCCTACGAGAAGAAAGTGCACAACGTGTGGGCCACTCACGCCTGCGTTCCCA
 CCGATCCTAATCCTCAAGAGATGGTGCTGAAGAACGTGACCGAGAACTTCAACATGTGGAAGAACGACATGG
 TGGACCAGATGCACGAGGACGTGATCAGCCTGTGGGACCAGAGCCTGAAGCCTTGCGTGAAGCTGACCCCTC
 TGTGTGTGACCCTGCACTGCAGCACCTACAACAACACCCACAACATCAGCAAGGGCATGAAGAACTGCAGCTT
 CAACGCCACCCAGAGCTGCGGGACAAGAAACAGAAAGTGAACGCCCTGTTCTACAAGCTGGACATCGTGCA
 GCTGAACAAGAACGGCCGGCAGTACCGGCTGATCAACTGCAACACCAGCGTGATCACCCAGGCCTGTCCTAA
 GGTGTCCTTCGATCCCATTCCTATCCACTACTGTGCCCCTGCCGGCTACGCCATCCTGAAGTGCAACAACAAGA
 CCTTCAACGGCACAGGCCCTGCAACAACGTGTCCACCGTGCAAGTGTACCCACGGCATCAAGCCTGTGGTGTG
 TACCCAGCTGCTGCTGAATGGATCTCTGGCCGAGGGCGAGATCATCATCAGAAGCGAGAACATCACCAACAA
 CGTCAAGACCATCATCGTCCACCTGAACGAGAGCGTGAAGATCGAGTGCACCCGGCCTAACAACAAAACCCG
 GACCAGCATCAGAATCGGCCCTGGCCAGACCTTTTACGCCACCGGACAAGTGTGCGGACATCAGAGAGGC
 CTAAGTGCACATCAACAAGAGCAAGTGGAAACGAGACACTGCAGCGGGTGTCCAAGAAGCTGAAAGAGTACTT
 CCCTCACAAGAATATCACCTTCAGCCTAGCTCTGGCGGGGACCTGGAAATCACCACACACAGCTTCAATTGTG
 GCGGGCAGTTCTTCTACTGCAATACCTCCAGCCTGTTCAACCGGACCTACATGGCCAACTCCACCGATATGGCC
 AACAGCACCGAGACAAACAGCACCAAGAACCATCACCATCCACTGCCGGATCAAGCAGATCATCAATATGTGGC
 AAGAAGTCGGCAGGGCTATGTACGCCCTCCTATCGCCGGCAACATCACCTGTATCAGCAATATCACCGGCCT
 GCTGCTCACCAGAGATGGCGGCAAGAACAACACCGAAACCTTCAGACCCGGCGGAGGCAACATGAAGGACA
 ATTGGAGAAGCGAGCTGTACAAGTACAAGGTGGTCAAGATCGAGCCCTGGGCGTCGCACCCACCAGGTGCA
 AAAGAAGAGTCGTGGAAGAAGGCGGAGGCGGAGAGCCGTTGGAATTGGAGCAGTGTTCTGGGCTTCTG
 GGAGCCGCGGATCTACAATGGGAGCTGCCAGCATGACCCTGACCGTGCAAGGCTAGAAATCTGCTGAGCGGC
 ATTGTGCAGCAGCAGAGCAACCTGCTGAGAGCCCTGAAGCTCAGCAGCACCTCCTGAAACTGACCGTGTGG
 GGAATCAAGCAGCTGCAGGCAAGAGTGCTGGCAGTGGAAGATACCTGCGGGACCAGCAGCTCCTCGGAAT
 CTGGGGATGTAGCGGCAAGCTGATCTGCTGCACCAACGTGCCCTGGAACAGCTCCTGGTCCAACAGAAACCT
 GAGCGAGATCTGGGACAACATGACCTGGCTGCAGTGGGACAAAGAGATCAGCAACTACACCCAGATCATCTA
 CGGACTGCTGGAAGAGAGCCAGAACCAGCAAGAGAAGAACGAGCAGGACCTGCTGGCCCTGGACTGATAAG
GATCC

>HV1301911 (CH505TF_V2.SET.OPT.N332_ch.SOSIPv4.1)

GTCGACGCCACCATGCTTATGGGATCTCTGCAGCCTCTGGCCACACTGTACCTGCTGGGAATGCTGGTGGCTT
 CTGTGCTGGCCGCCGAGAATCTGTGGGTACAGTGTACTATGGCGTGCCCGTGTGGAAAGAGGCCAAGACCA
 CACTGTTCTGTGCCAGCGACGCCAAGGCCTACGAGAAGAAAGTGCACAACGTGTGGGCCACTCACGCCTGCG
 TTCCACCGATCCTAATCCTCAAGAGATGGTGCTGAAGAACGTGACCGAGAACTTCAACATGTGGAAGAACG
 ACATGGTGGACCAGATGCACGAGGACGTGATCAGCCTGTGGGACCAGAGCCTGAAGCCTTGCGTGAAGCTG
 ACCCCTCTGTGTGTGACCCTGCACTGCAGCACCTACAACAACACCCACAACATCAGCAAGGGCATGAAGAACT
 GCAGCTTCAACGCCACCCAGAGCTGCGGGACAAGAAACAGAAAGTGAACGCCCTGTTCTACAAGCTGGACA
 TCGTGCAGCTGAACAAGAACGGCCGGCAGTACCGGCTGATCAACTGCAACACCAGCGTGATCACCCAGGCCT
 GTCCTAAGGTGTCCTTCGATCCCATTCCTATCCACTACTGTGCCCCTGCCGGCTACGCCATCCTGAAGTGCAAC
 AACAAGACCTTCAACGGCACAGGCCCTGCAACAACGTGTCCACCGTGCAAGTGTACCCACGGCATCAAGCCTG
 TGGTGTCTACCCAGCTGCTGCTGAATGGATCTCTGGCCGAGGGCGAGATCATCATCAGAAGCGAGAACATCA
 CCAACAACGTCAAGACCATCATCGTCCACCTGAACGAGAGCGTGAAGATCGAGTGCACCCGGCCTAACAACA
 AAACCCGGACCAGCATCAGAATCGGCCCTGGCCAG

Figure 4A continued

ACCTTTTACGCCACCGGACAAGTGATCGGGCGACATCAGAGAGGCCTACTGCAACATCTCCAAGAGCAAGTGG
AACGAGACACTGCAGCGGGTGTCCAAGAAGCTGAAAGAGTACTTCCCTCACAAGAATATCACCTTCCAGCCTA
GCTCTGGCGGGCGACCTGGAAATCACCACACACAGCTTCAATTGTGGCGGGAGTTCTTCTACTGCAATACCTC
CAGCCTGTTCAACCGGACCTACATGGCCAACTCCACCGATATGGCCAACAGCACCGAGACAAACAGCACCA
ACCATCACCATCCACTGCCGGATCAAGCAGATCATCAATATGTGGCAAGAAGTCGGCAGGGCTATGTACGCCC
CTCCTATCGCCGGCAACATCACCTGTATCAGCAATATCACCGGCCTGCTGCTCACCAGAGATGGCGGCAAGAA
CAACACCGAAACCTTCAGACCCGGCGGAGGCAACATGAAGGACAATTGGAGAAGCGAGCTGTACAAGTACA
AGGTGGTCAAGATCGAGCCCCTGGGCGTCGCACCCACCAGGTGCAAAAGAAGAGTCGTCGGAAGAAGGCGG
AGGCGGAGAGCCGTTGGAATTGGAGCAGTGTTCTGGGCTTTCTGGGAGCCGCCGGATCTACAATGGGAGCT
GCCAGCATGACCCTGACCGTGCAGGCTAGAAATCTGCTGAGCGGCATTGTGCAGCAGCAGAGCAACCTGCTG
AGAGCCCCTGAAGCTCAGCAGCACCTCCTGAAACTGACCGTGTGGGGAATCAAGCAGCTGCAGGCAAGAGTG
CTGGCAGTGGAAAGATACCTGCGGGACCAGCAGCTCCTCGGAATCTGGGGATGTAGCGGCAAGCTGATCTGC
TGCACCAACGTGCCCTGGAACAGCTCCTGGTCCAACAGAAACCTGAGCGAGATCTGGGACAACATGACCTGG
CTGCAGTGGGACAAAGAGATCAGCAACTACACCCAGATCATCTACGGAATGCTGGAAGAGAGCCAGAACCAG
CAAGAGAAGAACGAGCAGGACCTGCTGGCCCTGGACTGATAAGGATCC

>HV1301912 (CH505TF_V2.UCA.OPT1_ch.SOSIPv4.1)

GTGACGCCACCATGCTATGGGATCTCTGCAGCCTCTGGCCACACTGTAGCTGCTGGGAATGCTGCTGGCTT
CTGTGCTGGCGGCCGAGAATCTGTGGGTCAACGTGTACTATGGCGTGCCCGTGTGGAAAGAGGCCAAGACCA
CACTGTTCTGTGCCAGCGACGCCAAGGCCTACGAGAAGAAAGTGCAACAGTGTGGGCCACTCACGCCTGCG
TTCACCGATCCTAATCCTCAAGAGATGGTGCTGAAGAACGTGACCGAGAAGTTCAACATGTGGAAGAACG
ACATGGTGGACCAGATGCACGAGGACGTGATCAGCCTGTGGGACCAGAGCCTGAAGCCTTGCGTGAAGCTG
ACCCCTCTGTGTGTGACCCTGCACTGCAGCACCTACAACAACACCCACAACATCAGCAAGGGCATGAAGAAGT
GCAGCTTCAATATGACCACCGAGCTGCGGGACAAGAAGCAGAAAGTCAACGCCCTGTTCTACAAGCTGGACA
TCGTGCAGCTGAACAAGAACGGCCGGCAGTACCGGCTGATCAACTGCAACACAAGCGCCATCACACAGGCTT
GCCCCAAGGTGTCCTTCGATCCCATTCCTATCCACTACTGTGCCCTGCCGGCTACGCCATCCTGAAGTGCAAC
AACAAGACCTTCAACGGCACAGGCCCTGCAACAACGTGTCCACCGTGCAGTGTACCCACGGCATCAAGCCAG
TGGTGTCTACCCAGCTGCTGCTGAATGGATCTCTGGCCGAGGGCGAGATCATCATCAGAAGCGAGAACATCA
CCAACAACGTCAAGACCATCATCGTCCACCTGAACGAGAGCGTGAAGATCGAGTGACCCGGCCTAACAACA
AAACCAGAAAGAGCATCCGGATCGGCCCTGGCCAGACCTTTTATGCTACTGGCGACGTGATCGGGCGACATCA
GAGAGGCCTACTGCAACATCAACAAGAGCAAGTGGAACGAGACACTGCAGCGGGTGTCCAAGAACTGAAA
GAGTACTTCCCGCACAAGAATATCACCTTCCAGCCTAGCTCTGGCGGGCAGCTGGAAATCACCACACACAGCT
TTAACTGTGGCGGGCAGTTCTTCTACTGCAATACCAGCAGCCTGTTCAACCGGACCTACATGGCCAACTCCACC
GATATGGCCAACAGCACCGAGACAAACAGCACCCGGACCATCACCATCCACTGCCGGATCAAGCAGATCATC
AATATGTGGCAAGAAGTCGGCAGGGCTATGTACGCCCTCCTATCGCCGGCAACATCACCTGTATCAGCAATA
TCACCGGCCTGCTGCTCACCAGAGATGGCGGCAAGAACAACACCGAAACCTTCAGACCCGGCGGAGGCAACA
TGAAGGACAATTGGAGAAGCGAGCTGTACAAGTACAAGGTGGTCAAGATCGAGCCCCTGGGCGTCGCACCC
ACCAGGTGCAAAAGAAGAGTCGTCGGAAGAAGGCGGAGGCGGAGAGCCGTTGGAATTGGAGCTGTGTTCT
GGGCTTTCTGGGAGCCGCCGGATCTACAATGGGAGCTGCCAGCATGACCCTGACCGTGC

Figure 4A continued

AGGCTAGAAATCTGCTGAGCGGCATTGTGCAGCAGCAGAGCAACCTGCTGAGAGCCCCTGAAGCTCAGCAGC
ACCTCCTGAAACTGACCGTGTGGGGAATCAAGCAGCTGCAGGCAAGAGTGCTGGCAGTGGAAAGATACCTGC
GGGACCAGCAGCTCCTCGGAATCTGGGGATGTAGCGGCAAGCTGATCTGCTGCACCAACGTGCCCTGGAACA
GCTCCTGGTCCAACAGAAACCTGAGCGAGATCTGGGACAACATGACCTGGCTGCAGTGGGACAAAGAGATCA
GCAACTACACCCAGATCATCTACGGACTGCTGGAAGAGAGCCAGAACCAGCAAGAGAAGAACGAGCAGGAC
CTGCTGGCCCTGGACTGATAAGGATCC

>HV1301913 (CH505TF_V2.UCA.OPT1.N332_ch.SOSIP.v4.1)

GTCGACGCCACCATGCTTATGGGATCTCTGCAGCCTCTGGCCACACTGTAACTGCTGGGAATGCTGGTGGCTT
CTGTGCTGGCCGCCGAGAATCTGTGGGTACAGTGTACTATGGCGTGCCCGTGTGGAAAGAGGCCAAGACCA
CACTGTTCTGTGCCAGCGACGCCAAGGCCTACGAGAAGAAAGTGCAACACGTGTGGGCCACTCACGCCTGCG
TTCCACCGATCCTAATCCTCAAGAGATGGTGCTGAAGAACGTGACCGAGAACTTCAACATGTGGAAGAACG
ACATGGTGGACCAGATGCACGAGGACGTGATCAGCCTGTGGGACCAGAGCCTGAAGCCTTGCGTGAAGCTG
ACCCCTCTGTGTGTGACCCTGCACTGCAGCACCTACAACAACACCCACAACATCAGCAAGGGCATGAAGAACT
GCAGCTTCAATATGACCACCGAGCTGCGGGACAAGAAGCAGAAAGTCAACGCCCTGTTCTACAAGCTGGACA
TCGTGCAGCTGAACAAGAACGGCCGGCAGTACCGGCTGATCAACTGCAACACAAGCGCCATCACACAGGCTT
GCCCCAAGGTGTCCTTCGATCCCATTCCTATCCACTACTGTGCCCTGCCGGCTACGCCATCCTGAAGTGCAAC
AACAAGACCTTCAACGGCACAGGCCCTGCAACAACGTGTCCACCGTGCAAGTGTACCCACGGCATCAAGCCAG
TGGTGTCTACCCAGCTGCTGCTGAATGGATCTCTGGCCGAGGGCGAGATCATCATCAGAAGCGAGAACATCA
CCAACAACGTCAAGACCATCATCGTCCACCTGAACGAGAGCGTGAAGATCGAGTGACCCCGGCCTAACAACA
AAACCAGAAAGAGCATCCGGATCGGCCCTGGCCAGACCTTTTATGCTACTGGCGACGTGATCGGCGACATCA
GAGAGGCCTACTGCAACATCTCCAAGAGCAAGTGGAACGAGACACTGCAGCGGGTGTCCAAGAACTGAAA
GAGTACTTCCCGCACAAGAATATCACCTTCAGCCTAGCTCTGGCGGCGACCTGGAAATCACCACACACAGCT
TTAACTGTGGCGGCGAGTTCTTCTACTGCAATACCAGCAGCCTGTTCAACCGGACCTACATGGCCAACTCCACC
GATATGGCCAACAGCACCGAGACAAACAGCACCCGGACCATCACCATCCACTGCCGGATCAAGCAGATCATC
AATATGTGGCAAGAAGTCGGCAGGGCTATGTACGCCCCCTCCTATCGCCGGCAACATCACCTGTATCAGCAATA
TCACCGGCCTGCTGCTCACCAGAGATGGCGGCAAGAACAACACCGAAACCTTCAGACCCGGCGGAGGCAACA
TGAAGGACAATTGGAGAAGCGAGCTGTACAAGTACAAGGTGGTCAAGATCGAGCCCCTGGGCGTCGCACCC
ACCAGGTGCAAAAGAAGAGTCGTGGAAGAAGGCGGAGGCGGAGAGCCGTTGGAATTGGAGCTGTGTTCT
GGGCTTTCTGGGAGCCGCCGGATCTACAATGGGAGCTGCCAGCATGACCCTGACCGTGCAGGCTAGAAATCT
GCTGAGCGGCATTGTGCAGCAGCAGAGCAACCTGCTGAGAGCCCCTGAAGCTCAGCAGCACCTCCTGAAACT
GACCGTGTGGGGAATCAAGCAGCTGCAGGCAAGAGTGCTGGCAGTGGAAAGATACCTGCGGGACCAGCAGC
TCCTCGGAATCTGGGGATGTAGCGGCAAGCTGATCTGCTGCACCAACGTGCCCTGGAACAGCTCCTGGTCCAA
CAGAAACCTGAGCGAGATCTGGGACAACATGACCTGGCTGCAGTGGGACAAAGAGATCAGCAACTACACCCA
GATCATCTACGGACTGCTGGAAGAGAGCCAGAACCAGCAAGAGAAGAACGAGCAGGACCTGCTGGCCCTGG
ACTGATAAGGATCC

Figure 4A continued

Amino Acid sequences

>HV1301908 (CH505TF_V2.UCA.OPT1.gp41mut_ch.SOSIP.v4.1)

VDATMPMGSLQPLATLYLLGMLVASVLAAENLWVTVYYGVPVWKEAKTTLFCASDAKAYEKKVHNVWATHACV
 PTDPNPQEMVLKNVTENFNMWKNDMVDQM HEDVISLWDQSLKPCVKLTPLCVTLHCSTYNNTHNISKGMKNC
 SFNMTTEL RDKKQKVNALFYKLDIVQLNKNGRQYRLINCNTSAITQACPKVSFDPIPIHYCAPAGYAILKCNKTFNG
 TGPCNNVSTVQCTHGIKPVVSTQLLLNGSLAEGEIIIRSENITNNVKTIIVHLNESVKIECTRPNNKTRKSIRIGPGQTFY
 ATGDVIGDIREAYCNINKSKWNETLQRVSKKLKEYFPHKNITFQPSSGGDLEITTHSFNCGGEFFYCNTSSLFNRTYM
 ANSTD MANSTETNSTRITIHCRKQIINMWQEVGRAMYAPPIAGNITCISNITGLLLTRDGGKNNNTETFRPGGGN
 MKDNWRSELYKYKVVKIEPLGVAPTRCKRRRVVGRRRRRRAVGIGAVFLGFLGAAGSTMGAASITLTVQARNLLSGI
 VQQQSNLLRAPEAQQHLLKLTWVGKQLQARVLAVERYLKDQQLGIWGC SGKLICTNVPWNSSWSNRNLSEI
 WDNMTWLQWDKEISNYTQIYRLLEESQNQQEKNEQDLLALD

>HV1301909 (CH505TF_V2.UCA.OPT1.N332.gp41mut_ch.SOSIP.v4.1)

VDATMPMGSLQPLATLYLLGMLVASVLAAENLWVTVYYGVPVWKEAKTTLFCASDAKAYEKKVHNVWATHACV
 PTDPNPQEMVLKNVTENFNMWKNDMVDQM HEDVISLWDQSLKPCVKLTPLCVTLHCSTYNNTHNISKGMKNC
 SFNMTTEL RDKKQKVNALFYKLDIVQLNKNGRQYRLINCNTSAITQACPKVSFDPIPIHYCAPAGYAILKCNKTFNG
 TGPCNNVSTVQCTHGIKPVVSTQLLLNGSLAEGEIIIRSENITNNVKTIIVHLNESVKIECTRPNNKTRKSIRIGPGQTFY
 ATGDVIGDIREAYCNINKSKWNETLQRVSKKLKEYFPHKNITFQPSSGGDLEITTHSFNCGGEFFYCNTSSLFNRTYM
 ANSTD MANSTETNSTRITIHCRKQIINMWQEVGRAMYAPPIAGNITCISNITGLLLTRDGGKNNNTETFRPGGGN
 MKDNWRSELYKYKVVKIEPLGVAPTRCKRRRVVGRRRRRRAVGIGAVFLGFLGAAGSTMGAASITLTVQARNLLSGI
 VQQQSNLLRAPEAQQHLLKLTWVGKQLQARVLAVERYLKDQQLGIWGC SGKLICTNVPWNSSWSNRNLSEI
 WDNMTWLQWDKEISNYTQIYRLLEESQNQQEKNEQDLLALD

>HV1301910 (CH505TF_V2.SET.OPT_ch.SOSIPv4.1)

VDATMPMGSLQPLATLYLLGMLVASVLAAENLWVTVYYGVPVWKEAKTTLFCASDAKAYEKKVHNVWATHACV
 PTDPNPQEMVLKNVTENFNMWKNDMVDQM HEDVISLWDQSLKPCVKLTPLCVTLHCSTYNNTHNISKGMKNC
 SFNATTEL RDKKQKVNALFYKLDIVQLNKNGRQYRLINCNTSVITQACPKVSFDPIPIHYCAPAGYAILKCNKTFNG
 TGPCNNVSTVQCTHGIKPVVSTQLLLNGSLAEGEIIIRSENITNNVKTIIVHLNESVKIECTRPNNKTRTSIRIGPGQTFY
 ATGQVIGDIREAYCNINKSKWNETLQRVSKKLKEYFPHKNITFQPSSGGDLEITTHSFNCGGEFFYCNTSSLFNRTYM
 ANSTD MANSTETNSTRITIHCRKQIINMWQEVGRAMYAPPIAGNITCISNITGLLLTRDGGKNNNTETFRPGGGN
 MKDNWRSELYKYKVVKIEPLGVAPTRCKRRRVVGRRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLS
 GIVQQQSNLLRAPEAQQHLLKLTWVGKQLQARVLAVERYLRDQQLGIWGC SGKLICTNVPWNSSWSNRNLS
 EIWDNMTWLQWDKEISNYTQIYGLLEESQNQQEKNEQDLLALD

Figure 4B

>HV1301911 (CH505TF_V2.SET.OPT.N332_ch.SOSIPv4.1)

VDATMPMGSLOPLATLYLLGMLVASVLAAENLWVTVYYGVPVWKEAKTTLFCASDAKAYEKKVHNVWATHACV
 PTDPNPQEMVLKNVTENFNMWKNDMVDQM HEDVISLWDQSLKPCVKLTPLCVTLHCSTYNNTHNISKGMKNC
 SFNATTEL RDKKQKVNALFYKLDIVQLNKNGRQYRLINCNTSVITQACPKVSFDPIPIHYCAPAGYAILKCNKTFNG
 TGPCNNVSTVQCTHGIKPVVSTQLLLNGSLAEGEIIIRSENITNNVKTIVHLNESVKIECTRPNNKTRTSIRIGPGQTFY
 ATGQVIGDIREAYCNISKSKWNETLQRVSKKLKEYFPHKNITFQPSSGGDLEITTHSFNCGGEFFYCNTSSLFNRTYM
 ANSTD MANSTETNSTRITIHCRKQIINMWQEVGRAMYAPPIAGNITCISNITGLLLTRDGGKNN TETFRPGGGN
 MKDNWRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLS
 GIVQQQSNLLRAPEAQQHLLKLT VWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLS
 EIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALD

>HV1301912 (CH505TF_V2.UCA.OPT1_ch.SOSIPv4.1)

VDATMPMGSLOPLATLYLLGMLVASVLAAENLWVTVYYGVPVWKEAKTTLFCASDAKAYEKKVHNVWATHACV
 PTDPNPQEMVLKNVTENFNMWKNDMVDQM HEDVISLWDQSLKPCVKLTPLCVTLHCSTYNNTHNISKGMKNC
 SFNMTTEL RDKKQKVNALFYKLDIVQLNKNGRQYRLINCNTSAITQACPKVSFDPIPIHYCAPAGYAILKCNKTFNG
 TGPCNNVSTVQCTHGIKPVVSTQLLLNGSLAEGEIIIRSENITNNVKTIVHLNESVKIECTRPNNKTRKSIRIGPGQTFY
 ATGDVIGDIREAYCNINKSKWNETLQRVSKKLKEYFPHKNITFQPSSGGDLEITTHSFNCGGEFFYCNTSSLFNRTYM
 ANSTD MANSTETNSTRITIHCRKQIINMWQEVGRAMYAPPIAGNITCISNITGLLLTRDGGKNN TETFRPGGGN
 MKDNWRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLS
 GIVQQQSNLLRAPEAQQHLLKLT VWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLS
 EIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALD

>HV1301913 (CH505TF_V2.UCA.OPT1.N332_ch.SOSIP.v4.1)

VDATMPMGSLOPLATLYLLGMLVASVLAAENLWVTVYYGVPVWKEAKTTLFCASDAKAYEKKVHNVWATHACV
 PTDPNPQEMVLKNVTENFNMWKNDMVDQM HEDVISLWDQSLKPCVKLTPLCVTLHCSTYNNTHNISKGMKNC
 SFNMTTEL RDKKQKVNALFYKLDIVQLNKNGRQYRLINCNTSAITQACPKVSFDPIPIHYCAPAGYAILKCNKTFNG
 TGPCNNVSTVQCTHGIKPVVSTQLLLNGSLAEGEIIIRSENITNNVKTIVHLNESVKIECTRPNNKTRKSIRIGPGQTFY
 ATGDVIGDIREAYCNISKSKWNETLQRVSKKLKEYFPHKNITFQPSSGGDLEITTHSFNCGGEFFYCNTSSLFNRTYM
 ANSTD MANSTETNSTRITIHCRKQIINMWQEVGRAMYAPPIAGNITCISNITGLLLTRDGGKNN TETFRPGGGN
 MKDNWRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLS
 GIVQQQSNLLRAPEAQQHLLKLT VWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLS
 EIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALD

Figure 4B continued

>CH505TF chimera SOSIP.664.v4.1 C-SORTA (HV1301189_C_SORTAv2)

ATGCCCATGGGSCAGCCTGCAGCCCCCTGGCCACCCTGTACCTGCTGGGCATGCTGGTGGCTAGCGTGCTGGC
CGCCGAGAACCTGTGGGTGACCGTCTACTATGGCGTGCCCGTCTGGAAGGAAGCCAAAACCACACTGTTCT
GCGCTAGCGACGCTAAGGCATACGAGAAAAAAGTGCACAATGTCTGGGCTACTCATGCATGCGTGCCTAC
CGATCCAAATCCCCAGGAGATGGTGCTGAAGAACGTACAGAAAACCTTTAATATGTGGAAGAACGACATG
GTGGATCAGATGCACGAGGACGTGATCAGCCTGTGGGATCAGTCCCTGAAGCCATGCGTGAAACTGACTC
CCCTGTGCGTCACCCTGAACTGTACTAATGCCACCGCTTCCAACAGCTCCATCATTGAGGGGGATGAAGAACT
GTTCTTTCAATATCACTACCGAGCTGCGCGACAAGCGAGAAAAGAAAAATGCCCTGTTTTACAACTGGAC
ATCGTGCACTGGATGGCAACTCTAGTCAGTATAGACTGATTAAGTCAATACAAGCGTGATCACTCAGGC
ATGTCCAAAGGTCAGTTTTGATCCTATTCCAATCCACTACTGCGCACCCGCCGGATATGCTATCCTGAAGTG
TAACAACAAGACCTTCACCGGCACTGGGCCTTGCAACAACGTGAGCACCGTCCAGTGTACACATGGCATT
AGCCAGTGGTCAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCAGAGGGCGAAATCATTATCCGCAGCGA
GAACATCACAAATAATGTGAAGACTATCATCGTCCACCTGAACGAGAGCGTGAAGATTGAATGCACACGGC
CCAACAACAAGACCAGGACATCCATTCGCATCGGACCTGGCCAGTGGTTCTACGCTACTGGCCAGGTCATC
GGGGACATCAGAGAGGCCTATTGTAACATCAATGAGTCAAAGTGGAATGAAACTCTGCAGAGGGTGAGCA
AGAAACTGAAGGAATACTTCCCTCACAAAAACATCACCTTTCAGCCATCAAGCGGCGGGGACCTGGAGATT
ACAACTCATTCTTTCAATTGCGGAGGCGAATTCTTTTACTGTAACACCTCCTCTCTGTTTAATCGCACATATAT
GGCTAACAGTACTGATATGGCAAACCTCTACTGAGACCAATAGTACACGAACTATTACCATCCATTGCCGGAT
CAAGCAGATTATCAACATGTGGCAGGAAGTGGGGCGGGCCATGTATGCTCCCCCTATTGCAGGAAATATTA
CCTGTATCAGCAACATTACCGGCCTGCTGCTGACAAGAGACGGGGGAAAGAACAATACAGAGACTTTTAG
GCCTGGCGGGGGAAACATGAAAGATAATTGGCGCTCCGAGCTGTACAAGTATAAAGTGGTCAAGATCGAA
CCACTGGGAGTGGCACCTACCCGATGTAAACGGAGAGTGGTCGGAAGGCGCCGACGGAGAAGGGGCAGTG
GGAATCGGAGCCGTCTTCCTGGGCTTTCTGGGAGCAGCTGGCAGCACAATGGGAGCAGCCTCTATGACCCT
GACAGTGCAGGCTCGAAATCTGCTGAGTGGGATCGTGCAGCAGCAGTCAAACCTGCTGCGAGCACCAGAG
GCACAGCAGCATCTGCTGAAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCAGAGTGCTGGCTGTCTG
AACGGTACCTGAGAGATCAGCAGCTGCTGGGAATCTGGGGATGCAGCGGAAAGCTGATTTGCTGTACAAA
CGTGCCCTGGAATAGTTCATGGTCAAACAGGAATCTGAGCGAGATCTGGGACAATATGACCTGGCTGCAG
TGGGATAAGGAAATCAGTAACTACACACAGATCATCTATGGCCTGCTGGAGGAATCACAGAACCAGCAGG
AGAAAAATGAACAGGACCTGCTGGCCCTGGATCTGCCTAGCACCGGATGATGA

Figure 5A

>CH505TF.6R.SOSIP.664.v4.1_N_SORTA (HV1301189_N-SORTA)

[illegible]

Figure 5B continued

>CH505M5chim.6R.SOSIP.664v4.1_N_SORTA

MEMGSLQPLATLYLLGMLVAQVLAAGGGGGGGGGGGGGGAENLWVTVYYGVVPVWKEAKTTLEFCASD
AKAYEKKVHNWATHACVPTDPNPQEMVLKNVTENFNMWKNDMVDQMHEVDVISLWDQSLKPCVKLT
PLCVTLNCTNATASNSSIIIEGMKNCSFNITTELRLDKREKKNALFYKLDIVQLDGNSSQYRLINCNT
SVITQACPKVSFDPIPIHYCAPAGYAILKCNNKFTFTGTGPCNNVSTVQCTHGIKFPVVSTQLLLNGS
LAEGEIIIRSENITKNVKTIIIVHLNESVKIECTRPNNKTRTSIRIGPGQWIFYATGQVIGDIREAYC
NINESKWNETLQRVSKKLKEYFPHKNITFQPSSGGDLEITTHSFNCGGEFFYCNTSSLENRTYMAN
STDMANSTETNSTRITITIHCRIKQIINMWQEVGRAMYAPPIAGNITCISNITGLLLTRDGGKNNT
TFRPGGGNMKDNWRSELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRAVGIGAVFLGFLGAAGSTM
GAASMTLTVQARNLLSGIVQQQSNLLRAPEAQQHLLKLTVWGIKQLQARVLAVERYLRDQQLLGIW
GCSGKLI CCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLA
LD*

>CH505M5chim.6R.SOSIP.664v4.1_C_SORTA

ATGCCCATGGGTTCCTGAGGCGCTGGCCACCTGTACCTGCTGGGCATGCTGGTGGCCTCGGIG
CTGGCGGCCGAGAACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCCAAGACC
ACCCTGTTCTGCGCCTCCGACGCCAAGGCCTACGAGAAGAAGGTGCACAACGTGTGGGCCACCCAC
GCCTGCGTGCCCAACGACCCCAACCCCAAGGAGATGGTGCTGAAGAACGTGACCGAGAACTTCAAC
ATGTGGAAGAACGACATGGTGGACCAGATGCACGAGGACGTGATCTCCCTGTGGGACCAGTCCCTG
AAGCCCTGCGTGAAAGCTGACCCCCCTGTGCGTGACCCCTGAACTGCACCAACGCCACCGCCTCCAAC
TCCTCCATCATCGAGGGCATGAAGAACTGCTCCTTCAACATCACCACCGAGCTGCGCGACAAGCGC
GAGAAGAAGAACGCCCTGTTCTACAAGCTGGACATCGTGCAGCTGGACGGCAACTCCTCCCAGTAC
CGCCTGATCAACTGCAACACCTCCGTGATCACCCAGGCCTGCCCCAAGGTGTCCTTCGACCCCATC
CCCATCCACTACTGCGCCCCCGCCGGCTACGCCATCCTGAAGTGCAACAACAAGACCTTCACCGGC
ACCGGCCCCCTGCAACAACGTGTCCACCGTGCAGTGCACCCACGGCATCAAGCCCGTGGTGTCCACC
CAGCTGCTGCTGAACGGCTCCCTGGCCGAGGGCGAGATCATCATCCGCTCCGAGAACATCACCAAG
AACGTGAAGACCATCATCGTGCACCTGAACGAGTCCGTGAAGATCGAGTGCACCCGCCCAACAAC
AAGACCCGCACCTCCATCCGCATCGGCCCCGGCCAGTGGTTCACGCCACCGGCCAGGTGATCGGC
GACATCCGCGAGGCCTACTGCAACATCAACGAGTCCAAGTGGAACGAGACCCTGCAGCGCGTGTCC
AAGAAGCTGAAGGAGTACTTCCCCACAAGAACATCACCTTCCAGCCCTCCTCCGGCGGGCGACCTG
GAGATCACCAACCACTCCTTCAACTGCGGCGGGCGAGTTCTTCTACTGCAACACCTCCTCCCTGTTT
AACCGCACCTACATGGCCAACCTCCACCGACATGGCCAACCTCCACCGAGACCAACTCCACCCGCACC
ATCACCATCCACTGCCGCATCAAGCAGATCATCAACATGTGGCAGGAGGTGGGCCGCGCCATGTAC
GCCCCCCCCATCGCCGGCAACATCACCTGCATCTCCAACATCACCGGCCTGCTGCTGACCCGCGAC
GGCGGCAAGAACAACACCGAGACCTTCCGCCCCGGCGGGCGGCAACATGAAGGACAACCTGGCGCTCC
GAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCCTGCAAGCGC
CGCGTGGTGGGCGCGCGCCGCGCGCGCGCGCGCGTGGGCATCGGCGCCGTGTTCTTGGGCTTCCTG
GGCGCCGCGGGCTCCACCATGGGCGCGCGCTCCATGACCCTGACCGTGCAGGCCCGCAACCTGCTG
TCGGGCATCGTGCAGCAGCAGTCCAACCTGCTGCGCGCCCCCGAGGCCAGCAGCACTGCTGAAG
CTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGTGGCCGTGGAGCGCTACCTGCGCGAC
CAGCAGCTGCTGGGCATCTGGGGCTGCTCCGGCAAGCTGATCTGCTGCACCAACGTGCCCTGGAAC
TCCTCCTGGTCCAACCGCAACCTGTCCGAGATCTGGGACAACATGACCTGGCTGCAGTGGGACAAG
GAGATCTCCAACCTACACCCAGATCATCTACGGCCTGCTGGAGGAGTCCCAGAACCAGCAGGAGAAG
AACGAGCAGGACCTGCTGGCCCTGGACCTGCCTAGCACCGGATAG

Figure 5B continued

>CH505M5chim.6R.SOSIP.664v4.1__C__SORTA

MEFGSLQPLATLYLLGMIYASVLAENLWVTVYYGVPVWKEAKTTLEFCASDAKAYEKKVHNWATH
ACVPTDPNPQEMVLKNVTENFNMWKNDMVDQMHEVDVLSLWDQSLKPCVKLTPLCVTLNCTNATASN
SSIIEGMKNCSFNITTELDRDKREKKNALFYKLDIVQLDGNSSQYRLINCNTSVITQACPKVSFDPI
PIHYCAPAGYAILKCNNKTFTGTGPCNNVSTVQCTHGIKPVVSTQLLNGSLAEGEIIIRSENITK
NVKTIIVHLNESVKIECTRPNNKTRTSIRIGPGQWFYATGQVIGDIREAYCNINESKWNETLQRVS
KKLKEYFPHKNITFQPSSGGDLEITTHSFNCGGEFFYCNTSSLENRTYMANSTDMANSTETNSTRT
ITIHCRIKQIINMWQEVGRAMYAPPIAGNITCISNITGLLLTRDGGKNNTETFRPGGGNMKDNWRS
ELYKYKVVKIEPLGVAPTRCKRRVVGRRRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLL
SGIVQQQSNLLRAPEAQQHLLKLTWVGKQLQARVLAVERYLRDQQLLGIWGCSGKLCCTNVPWN
SSWSNRNLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALDLPSTG*

Figure 5B continued

C-terminal modification of SOSIP trimer

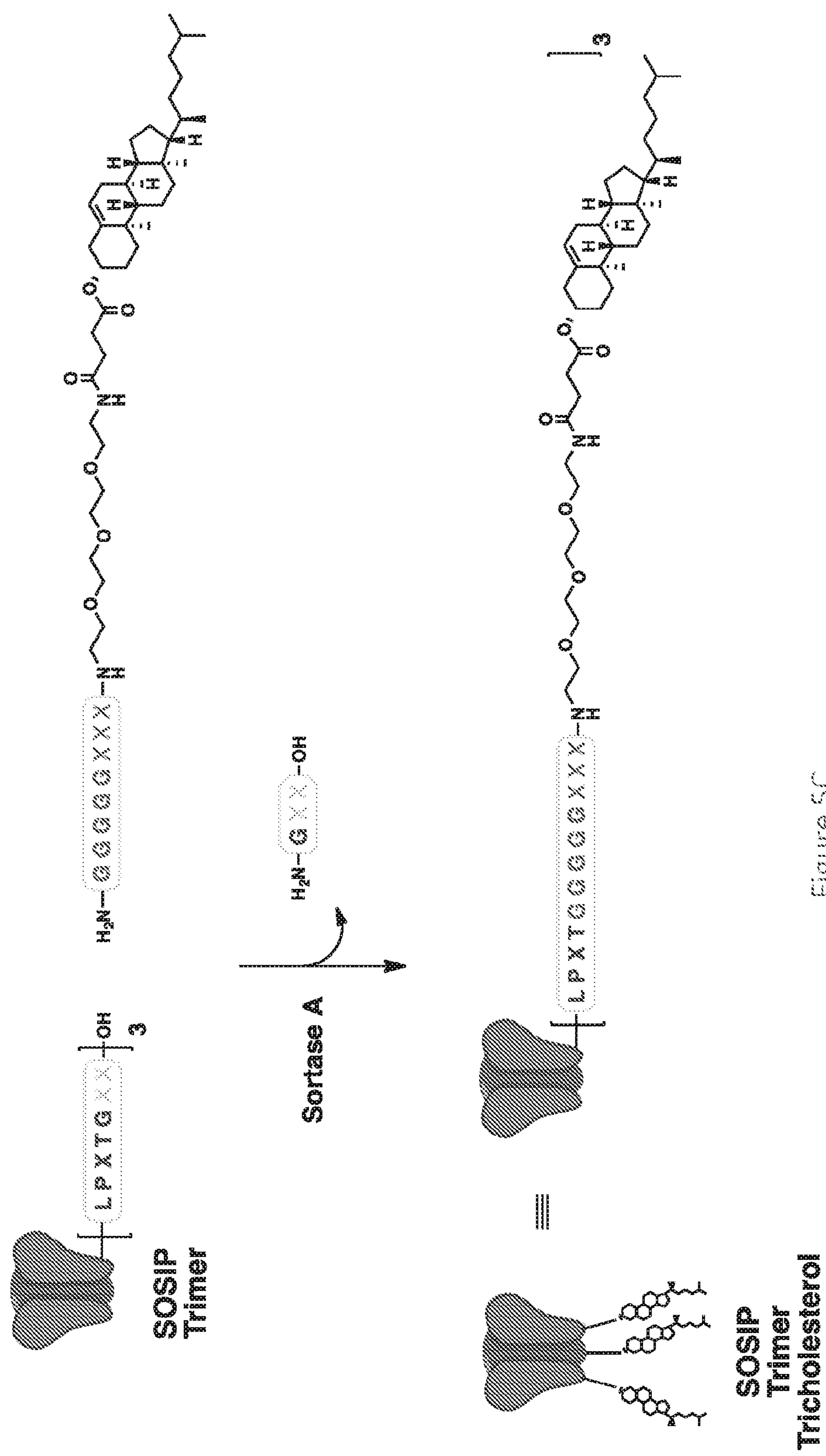


Figure 5C

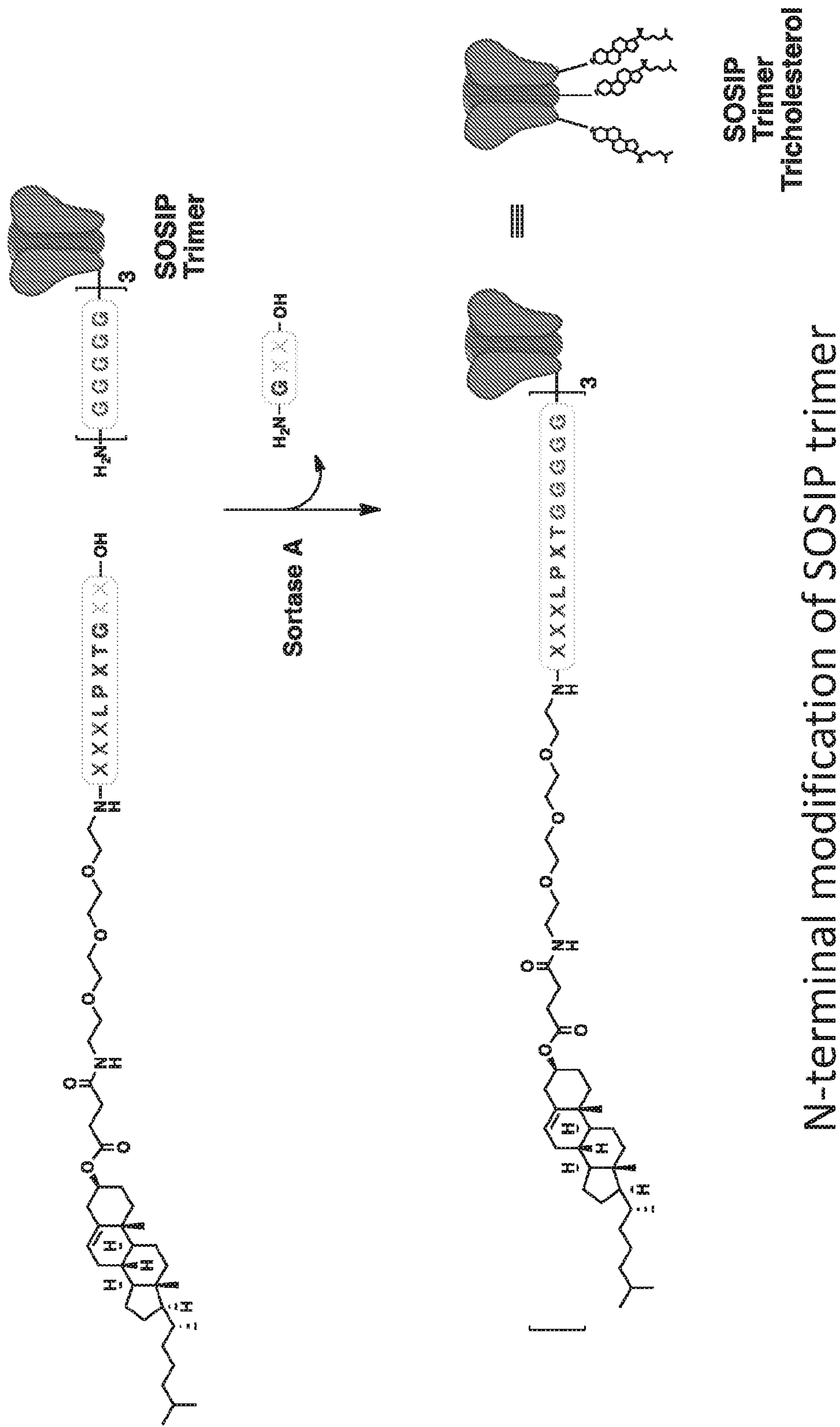
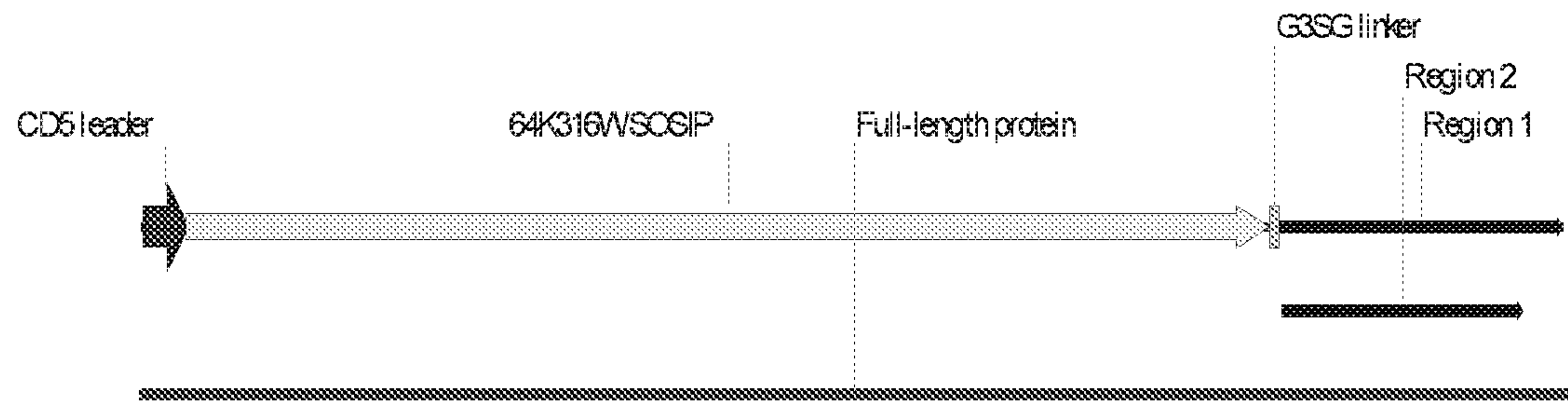


Figure 5D



CH505TFchimv4.1-ferritin
815 aa

>CH505TFchimv4.1-ferritin (signal peptide; ferritin, glycine linker)

MPMSLQPLA TLYLLGMLVA SVLAAENLWV TVYYGVPVWK EAKTTLFCAS
DAKAYEKKVH NVWATHACVP TDPNPQEMVL KNVTENENMW KNDMVDQMHE
DVISLWDQSL KPCVKLTPLC VTLNCTNATA SNSSIIEGMK NCSFNITTEL
RDKREKKNAL FYKLDIVQLD GNSSQYRLIN CNTSVITQAC PKVSFDFIPI
HYCAPAGYAI LKCNNKTFTG TGPCNNVSTV QCTHGIKPVV STQLLNLGSL
AEGEIIIRSE NITNNVKTHI VHLNESVKIE CTRPNNKTRT SIRIGPGQWF
YATGQVIGDI REAYCNINES KWNETHQVRS KKLKEYFPHK NITEQPSGG
DLEITHSFN CGGEFFYCNT SSLEFNRTYMA NSTDMANSTE TNSTRITITH
CRIKQIINMW QEVGRAMYAP PIAGNITCIS NITGLLLTRD GGKNNTEFER
PGGGNMKDNW RSELYKYKV V KIEPLGVAPT RCKRRVVGRR RRRRAVGIGA
VFLGFLGAAG STMGAASMTL TVQARNLLSG IVQQQSNLLR APEAQQHLLK
LTVWGIKQLQ ARVLAVERYL RDQQLLGIWG CSGKLICCTN VPWNSSWSNR
NLSEIWDNMT WLQWDKEISN YTQIIYGLLE ESQNQQEKNE QDLLALDGGG
SGDIIKLLNE QVNKEMNSSN LYMSMSSWCY THSLDGAGLF LFDHAAEEYE
HAKKLIIFLN ENNVFVQLTS ISAPEHKFEG LTQIFQKAYE HEQHISESIN
NIVDHAIKSK DHATFNFLQW YVAEQHEEEV LFKDILDKIE LIGNENHGLY
LADQYVKGIA KSRKS**

>CH505TFchimv4.1-ferritin (HV1301350)

ATGCCCATTGGCTGCTGCAAGCCCTGGCCACCCCTGTACCTGCTGGGCATGCTGGTGGCTC
CTGCTGGCCGCCGAGAACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGG
CCAAGACCACCCTGTTCTGCGCCTCCGACGCCAAGGCCTACGAGAAGAAGGTGCACAACGTG
TGGGCCACCCACGCCTGCGTGCCACCGACCCCAACCCCCAGGAGATGGTGCTGAAGAACGT
GACCGAGAACTTCAACATGTGGAAGAACGACATGGTGGACCAGATGCACGAGGACGTGATCT
CCCTGTGGGACCAGTCCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTGACCCTGAAC
TGCACCAACGCC

Figure 5E

ACCGCCTCCAACCTCCTCCATCATCGAGGGCATGAAGAACTGCTCCTTCAACATCACCACCGA
 GCTGCGCGACAAGCGCGAGAAGAAGAACGCCCTGTTCTACAAGCTGGACATCGTGCAGCTGG
 ACGGCAACTCCTCCCAGTACCGCCTGATCAACTGCAACACCTCCGTGATCACCAGGCCTGC
 CCCAAGGTGTCCTTCGACCCCATCCCCATCCACTACTGCGCCCCCGCCGGCTACGCCATCCT
 GAAGTGCAACAACAAGACCTTCACCGGCACCGGCCCTGCAACAACGTGTCCACCGTGCAGT
 GCACCCACGGCATCAAGCCCGTGGTGTCCACCCAGCTGCTGCTGAACGGCTCCCTGGCCGAG
 GGCGAGATCATCATCCGCTCCGAGAACATCACCAACAACGTGAAGACCATCATCGTGCACCT
 GAACGAGTCCGTGAAGATCGAGTGCACCCGCCCAACAACAAGACCCGCACCTCCATCCGCA
 TCGGCCCCGGCCAGTGGTTCTACGCCACCGGCCAGGTGATCGGCGACATCCGCGAGGCCTAC
 TGCAACATCAACGAGTCCAAGTGGAACGAGACCCTGCAGCGCGTGTCCAAGAAGCTGAAGGA
 GTACTTCCCCCACAAGAACATCACCTTCCAGCCCTCCTCCGGCGGCGACCTGGAGATCACCA
 CCCACTCCTTCAACTGCGGGCGGCGAGTTCTTCTACTGCAACACCTCCTCCCTGTTCAACCGC
 ACCTACATGGCCAACTCCACCGACATGGCCAACTCCACCGAGACCAACTCCACCCGCACCAT
 CACCATCCACTGCCGCATCAAGCAGATCATCAACATGTGGCAGGAGGTGGGCCGCGCCATGT
 ACGCCCCCCCCATCGCCGGCAACATCACCTGCATCTCCAACATCACCGGCCTGCTGCTGACC
 CGCGACGGCGGCAAGAACAACACCGAGACCTTCCGCCCCGGCGGCGGCAACATGAAGGACAA
 CTGGCGCTCCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCTGGGCGTGGCCCCCA
 CCCGCTGCAAGCGCCGCGTGGTGGGCGCCGCCGCCGCCGCCGCCGCGCGCGTGGGCATCGGCGCC
 GTGTTCTGTTGGGCTTCTGTTGGGCGCCGCCGCCGCTCCACCATGGGCGCCGCCCTCCATGACCCTGAC
 CGTGCAGGCCCGCAACCTGCTGTCCGGCATCGTGCAGCAGCAGTCCAACCTGCTGCGCGCCC
 CCGAGGCCAGCAGCACCTGCTGAAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGC
 GTGCTGGCCGTGGAGCGCTACCTGCGCGACCAGCAGCTGCTGGGCATCTGGGGCTGCTCCGG
 CAAGCTGATCTGCTGCACCAACGTGCCCTGGAACCTCCTCCTGGTCCAACCGCAACCTGTCCG
 AGATCTGGGACAACATGACCTGGCTGCAGTGGGACAAGGAGATCTCCAACCTACACCCAGATC
 ATCTACGGCCTGCTGGAGGAGTCCCAGAACCAGCAGGAGAAGAACGAGCAGGACCTGCTGGC
 CCTGGACGGCGGCGGCTCCGGCGACATCATCAAGCTGCTGAACGAGCAGGTGAACAAGGAGA
 TGAACCTCCTCCAACCTGTACATGTCCATGTCTCCTGGTGTACACCCACTCCCTGGACGGC
 GCCGGCCTGTTCTGTTTCGACCACGCCGCCGAGGAGTACGAGCACGCCAAGAAGCTGATCAT
 CTTCTGAACGAGAACAACGTGCCCGTGCAGCTGACCTCCATCTCCGCCCCGAGCACAAGT
 TCGAGGGCCTGACCCAGATCTTCCAGAAGGCCTACGAGCACGAGCAGCACATCTCCGAGTCC
 ATCAACAACATCGTGGACCACGCCATCAAGTCCAAGGACCACGCCACCTTCAACTTCTGCA
 GTGGTACGTGGCCGAGCAGCACGAGGAGGAGGTGCTGTTCAAGGACATCCTGGACAAGATCG
 AGCTGATCGGCAACGAGAACCACGGCCTGTACCTGGCCGACCAGTACGTGAAGGGCATCGCC
 AAGTCCCGCAAGTCCTGATAA

>CH505M5chimv4.1-ferritin (signal peptide; ferritin, glycine linker)

MPMSLQPLATLYLLGMLVASVLAENLWVTVYYGVPVWKEAKTTLFCASDAKAYEKKVH
 NVWATHACVPTDPNPQEMVLKNVTENFNMWKNMVDQMHEDEVISLWDQSLKPCVKLTPLC
 VTLNCTNATASNSSIIIEGMKNCSFNITTEL RDKREKKNALFYKLDIVQLDGNSSQYRLIN
 CNTSVITQACPKVSFDPIPIHYCAPAGYAILKCNNKTFTGTGPCNNVSTVQCTHGIKPVV

Figure 5E continued

STQLLLNGLAEGEIIIRSENITKNVKTIIIVHLNESVKIECTRPNNKTRTSIRIGPGQWF
 YATGQVIGDIREAYCNINESKWNETLQRVSKKLKEYFPHKNITFQPSSGGDLEITTHSFN
 CGGEFFYCNTSSLFNRTYMANSTDMANSTETNSTRITITIHCRKIQIINMWQEVGRAMYAP
 PIAGNITCISNITGLLLTRDGGKNNTEFRPGGGNMKDNWRSELYKYKVVKIEPLGVAPT
 RCKRRVVGRRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLR
 APEAQQHLLKLTVWGIKQLQARVLAVERYLRDQQLLGIWGC SGKLICCTNVPWNSSWSNR
 NLSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALDGGGSGDIIKLLNE
 QVNKEMNSSNLYMSMSWCYTHSLDGAGLFLFDHAAEHEYEHAKKLIIFLNENNVPVQLTS
 ISAPEHHKFEGLTQIFQKAYEHEQHISESINNIVDHAIKSKDHATFNFLQWYVAEQHEEEV
 LFKDILDKIELIGNENHGLYLADQYVKGIKSRKS**

>CH505M5chimv4.1-ferritin(HV1301349)

ATGCCCCATGGGTGCTGCAACCCCTGCCACCCCTGTACCTGCTGGGCATGCTGGTGGCCTG
 CCGTGGTGGCGCCGAGAACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGG
 CCAAGACCACCCCTGTTCTGCGCCTCCGACGCCAAGGCCTACGAGAAGAAGGTGCACAACGTG
 TGGGCCACCCACGCCTGCGTGCCACCGACCCCAACCCCCAGGAGATGGTGCTGAAGAACGT
 GACCGAGAACTTCAACATGTGGAAGAACGACATGGTGGACCAGATGCACGAGGACGTGATCT
 CCCTGTGGGACCAGTCCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTGACCCTGAAC
 TGCACCAACGCCACCGCCTCCAACCTCCTCCATCATCGAGGGCATGAAGAACTGCTCCTTCAA
 CATCACCACCGAGCTGCGCGACAAGCGCGAGAAGAAGAACGCCCTGTTCTACAAGCTGGACA
 TCGTGCACTGGACGGCAACTCCTCCCAGTACCGCCTGATCAACTGCAACACCTCCGTGATC
 ACCCAGGCCTGCCCCAAGGTGTCTTCGACCCCATCCCATCCACTACTGCGCCCCCGCCGG
 CTACGCCATCCTGAAGTGCAACAACAAGACCTTCACCGGCACCGGCCCTGCAACAACGTGT
 CCACCGTGCAGTGCACCCACGGCATCAAGCCCGTGGTGTCCACCCAGCTGCTGCTGAACGGC
 TCCCTGGCCGAGGGCGAGATCATCATCCGCTCCGAGAACATCACCAAGAACGTGAAGACCAT
 CATCGTGCACTGAACGAGTCCGTGAAGATCGAGTGCACCCGCCCCAACAACAAGACCCGCA
 CCTCCATCCGCATCGGCCCCCGGCCAGTGGTTCTACGCCACCGGCCAGGTGATCGGCGACATC
 CGCGAGGCCTACTGCAACATCAACGAGTCCAAGTGAACGAGACCCTGCAGCGCGTGTCCAA
 GAAGCTGAAGGAGTACTTCCCCCACAAGAACATCACCTTCCAGCCCTCCTCCGGCGGGCGACC
 TGGAGATCACCACCCACTCCTTCAACTGCGGCGGGCGAGTTCTTCTACTGCAACACCTCCTCC
 CTGTTCAACCGCACCTACATGGCCAACTCCACCGACATGGCCAACTCCACCGAGACCAACTC
 CACCCGCACCATCACCATCCACTGCCGCATCAAGCAGATCATCAACATGTGGCAGGAGGTGG
 GCCGCGCCATGTACGCCCCCCCCCATCGCCGGCAACATCACCTGCATCTCCAACATCACCGGC
 CTGCTGCTGACCCGCGACGGCGGCAAGAACAACACCGAGACCTTCCGCCCCGGCGGGCGGCAA
 CATGAAGGACAACTGGCGCTCCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGG
 GCGTGGCCCCCACCCTGCAAGCGCCGCGTGGTGGGCGCGCCGCGCCGCGCGCGCGCGT
 GGCATCGGCGCCGTGTTCTTGGGCTTCCTGGGCGCGCGGCTCCACCATGGGCGCGCGCTC
 CATGACCCTGACCGTGCAGGCCCCGCAACCTGCTGTCCGGCATCGTGCAGCAGCAGTCCAACC
 TGCTGCGCGCCCCGAGGCCAGCAGCACCTGCTGAAGCTGACCGTGTGGGG

Figure 5E continued

CATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGCGCGACCAGCAGCTGC
TGGGCATCTGGGGCTGCTCCGGCAAGCTGATCTGCTGCACCAACGTGCCCTGGAATCCTCC
TGGTCCAACCGCAACCTGTCCGAGATCTGGGACAACATGACCTGGCTGCAGTGGGACAAGGA
GATCTCCAATAACACCCAGATCATCTACGGCCTGCTGGAGGAGTCCCAGAACCAGCAGGAGA
AGAACGAGCAGGACCTGCTGGCCCTGGACGGCGGGCGGCTCCGGCGACATCATCAAGCTGCTG
AACGAGCAGGTGAACAAGGAGATGAACTCCTCCAACCTGTACATGTCCATGTCCTCCTGGTG
CTACACCCACTCCCTGGACGGCGCCGGCCTGTTCTGTTTCGACCACGCCGCCGAGGAGTACG
AGCACGCCAAGAAGCTGATCATCTTCCTGAACGAGAACAACGTGCCCCGTGCAGCTGACCTCC
ATCTCCGCCCCCGAGCACAAGTTCGAGGGCCTGACCCAGATCTTCCAGAAGGCCTACGAGCA
CGAGCAGCACATCTCCGAGTCCATCAACAACATCGTGGACCACGCCATCAAGTCCAAGGACC
ACGCCACCTTCAACTTCCTGCAGTGGTACGTGGCCGAGCAGCACGAGGAGGAGGTGCTGTTC
AAGGACATCCTGGACAAGATCGAGCTGATCGGCAACGAGAACCACGGCCTGTACCTGGCCGA
CCAGTACGTGAAGGGCATCGCCAAGTCCCGCAAGTCCTAGTAA

Figure 5E continued

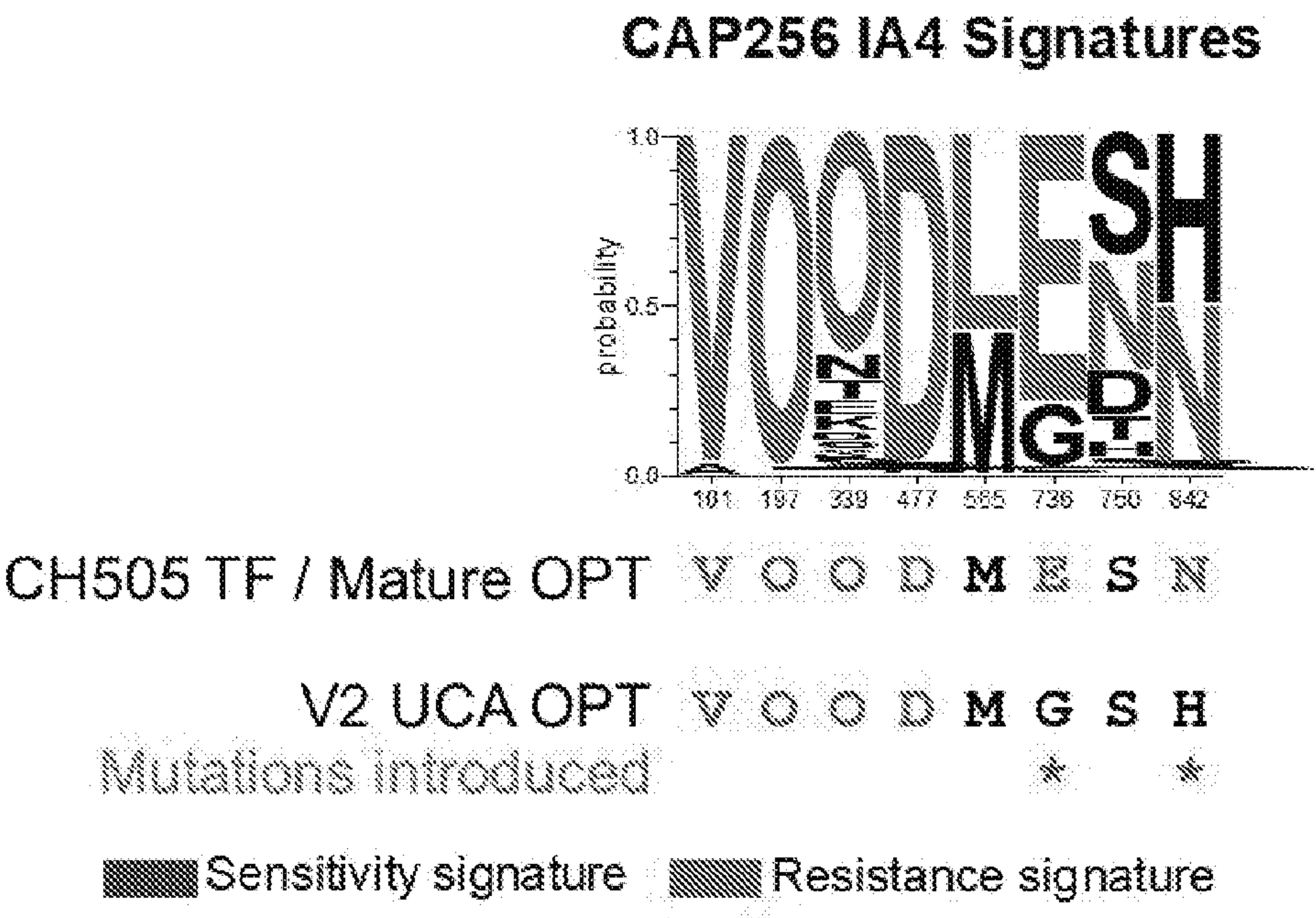


Figure 6

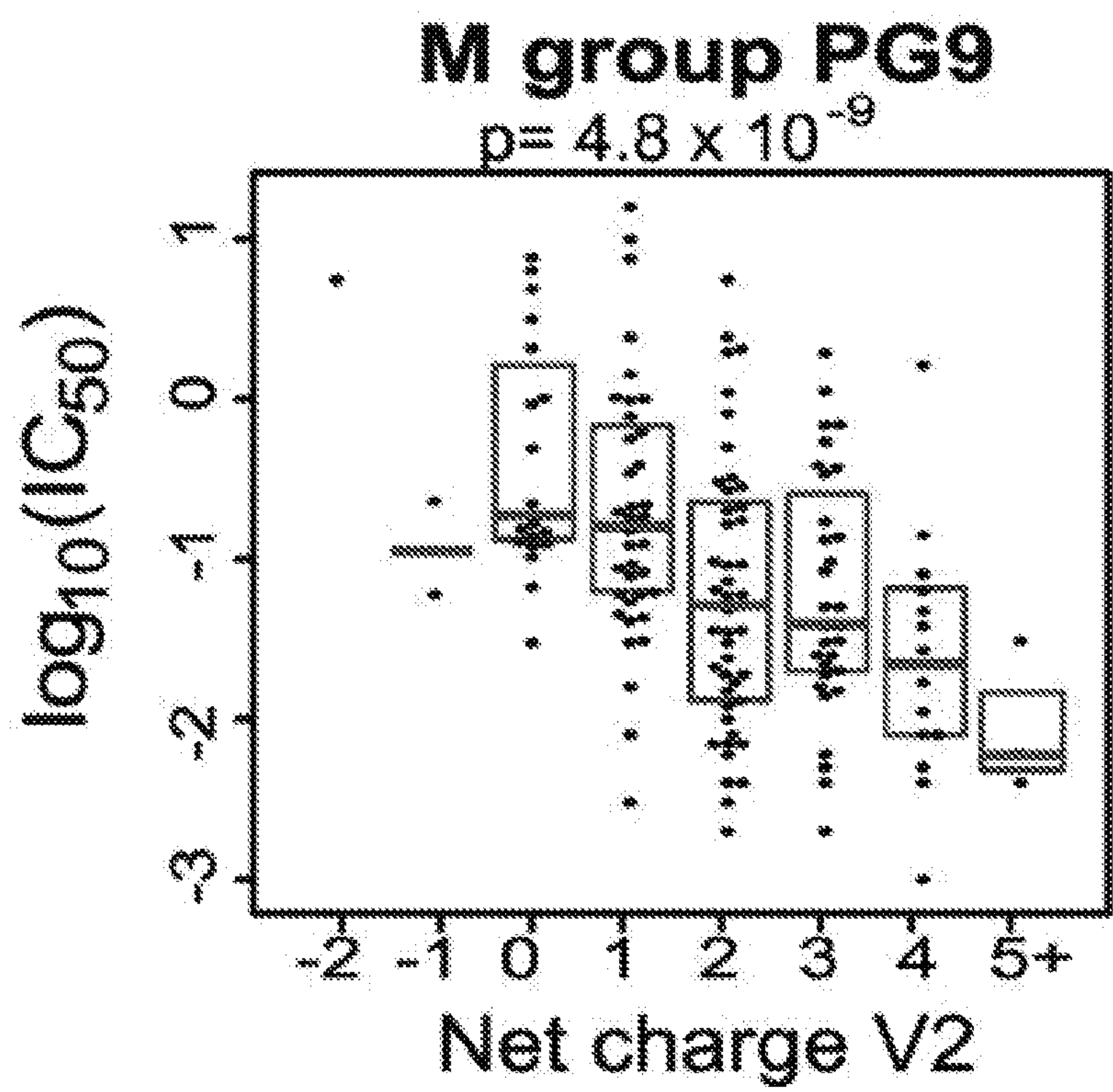


Figure 7

	<u>HYP V1:</u>	<u>HYP V2:</u>
CH505.TF	TNATASNSSIIE	DGNSSQ
ZM233.6	STYNNTHNISK	TNSSNTTN
T250-4	QAFNSSSHTNSSIAMQ	NKNGRQ
<u>Designed constructs</u>	<u>STYNNTHNISK</u>	<u>NKNGRQ</u>

Figure 8A

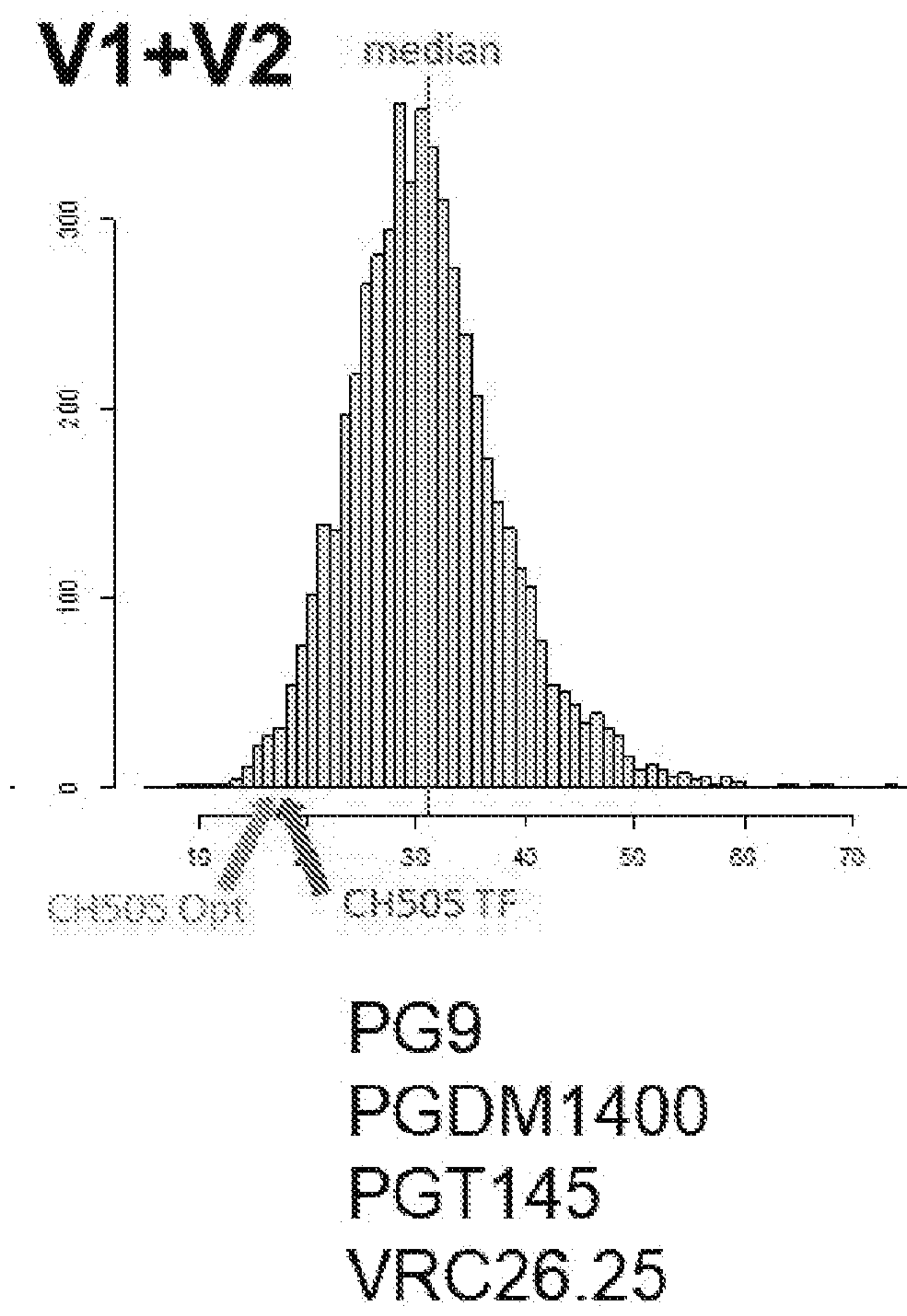


Figure 8B

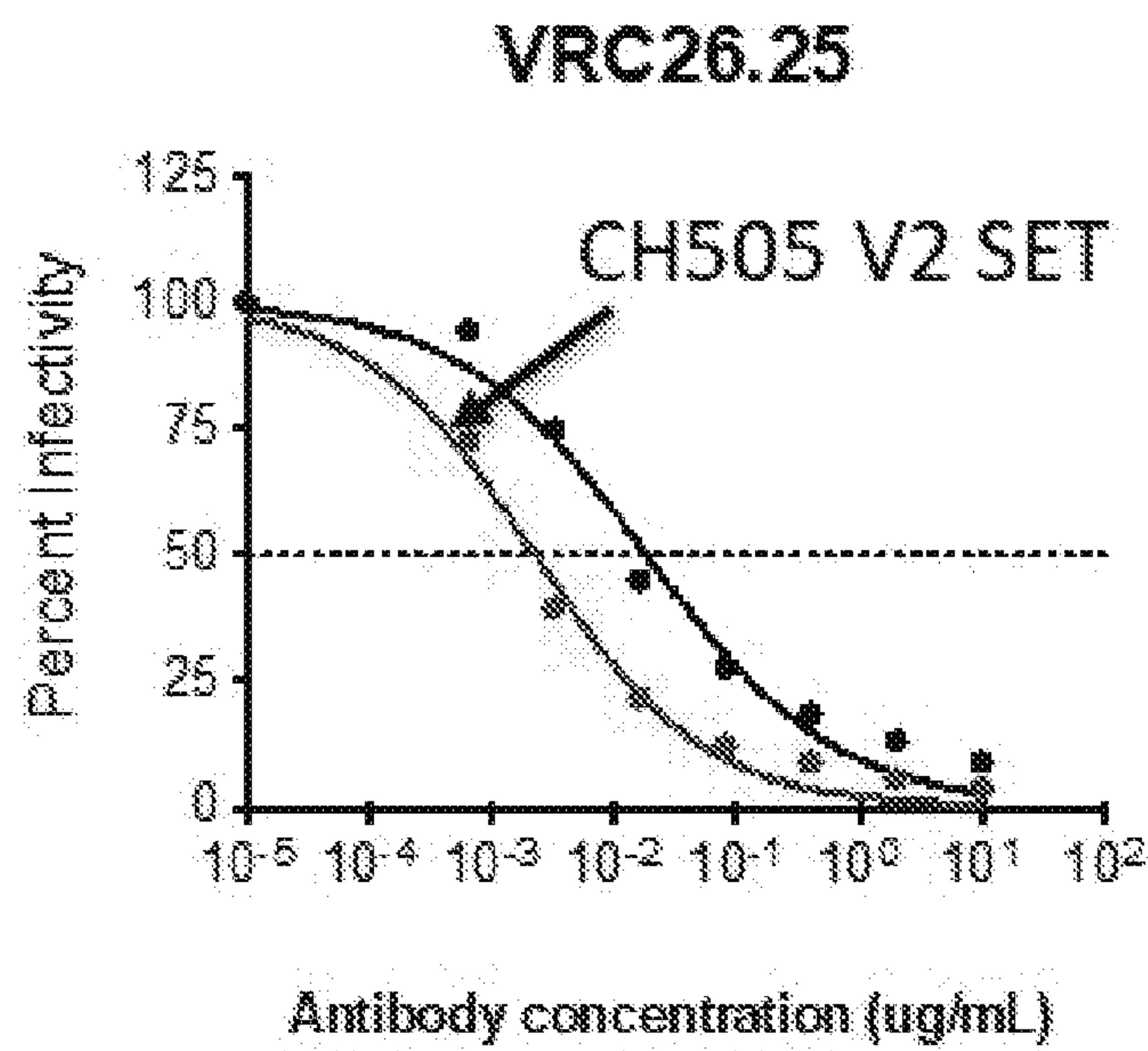


Figure 9A

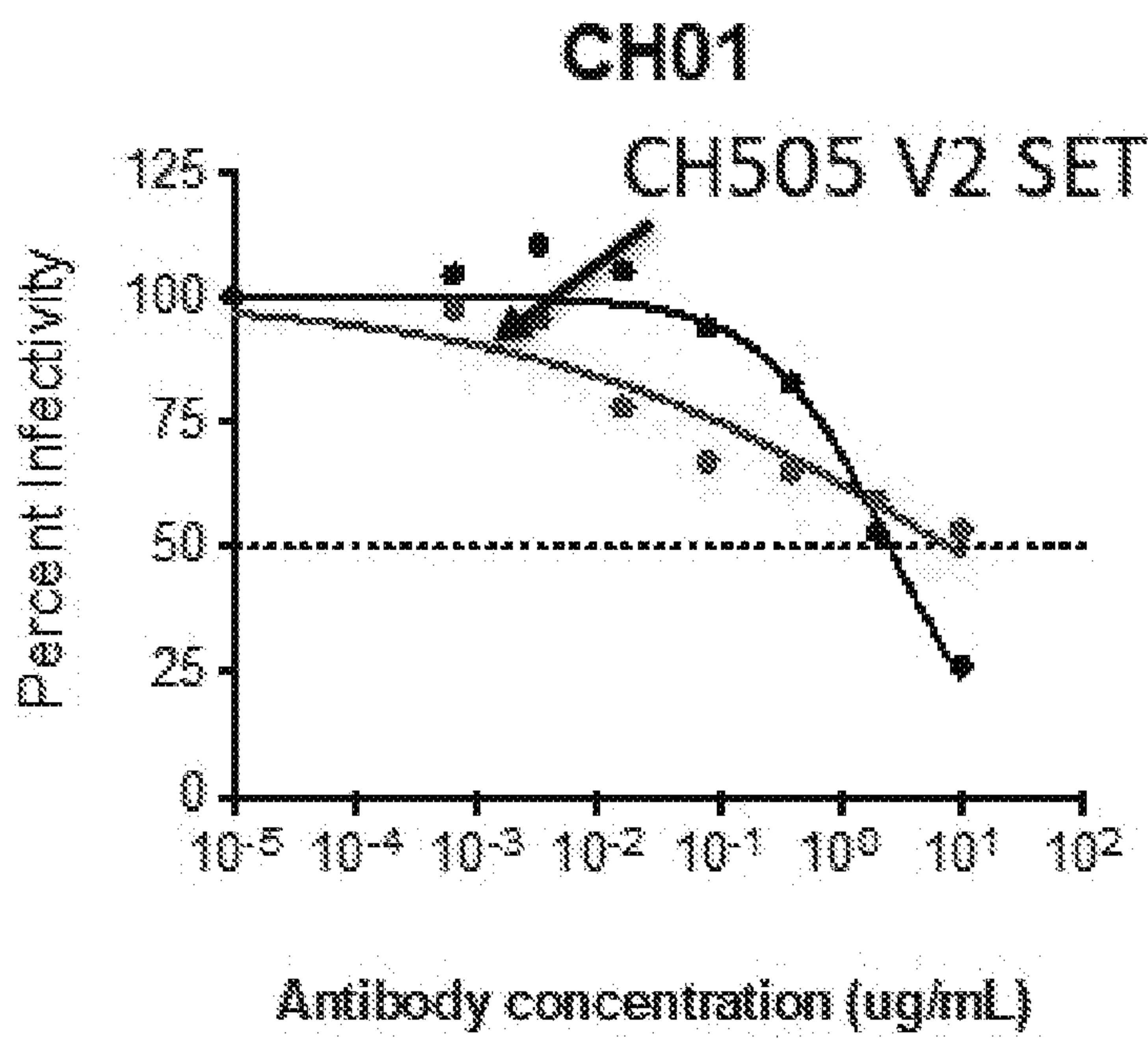


Figure 9B

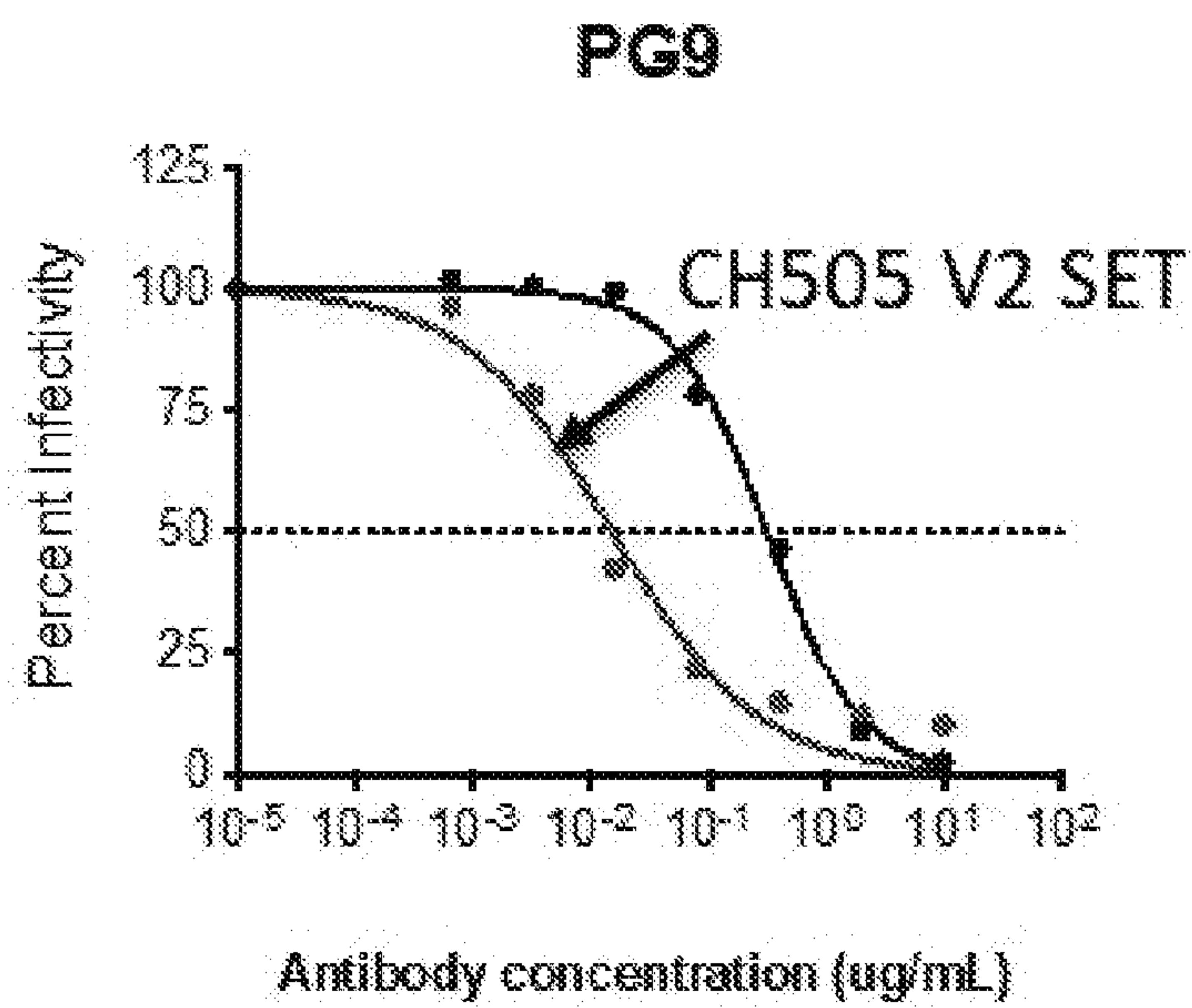


Figure 9C

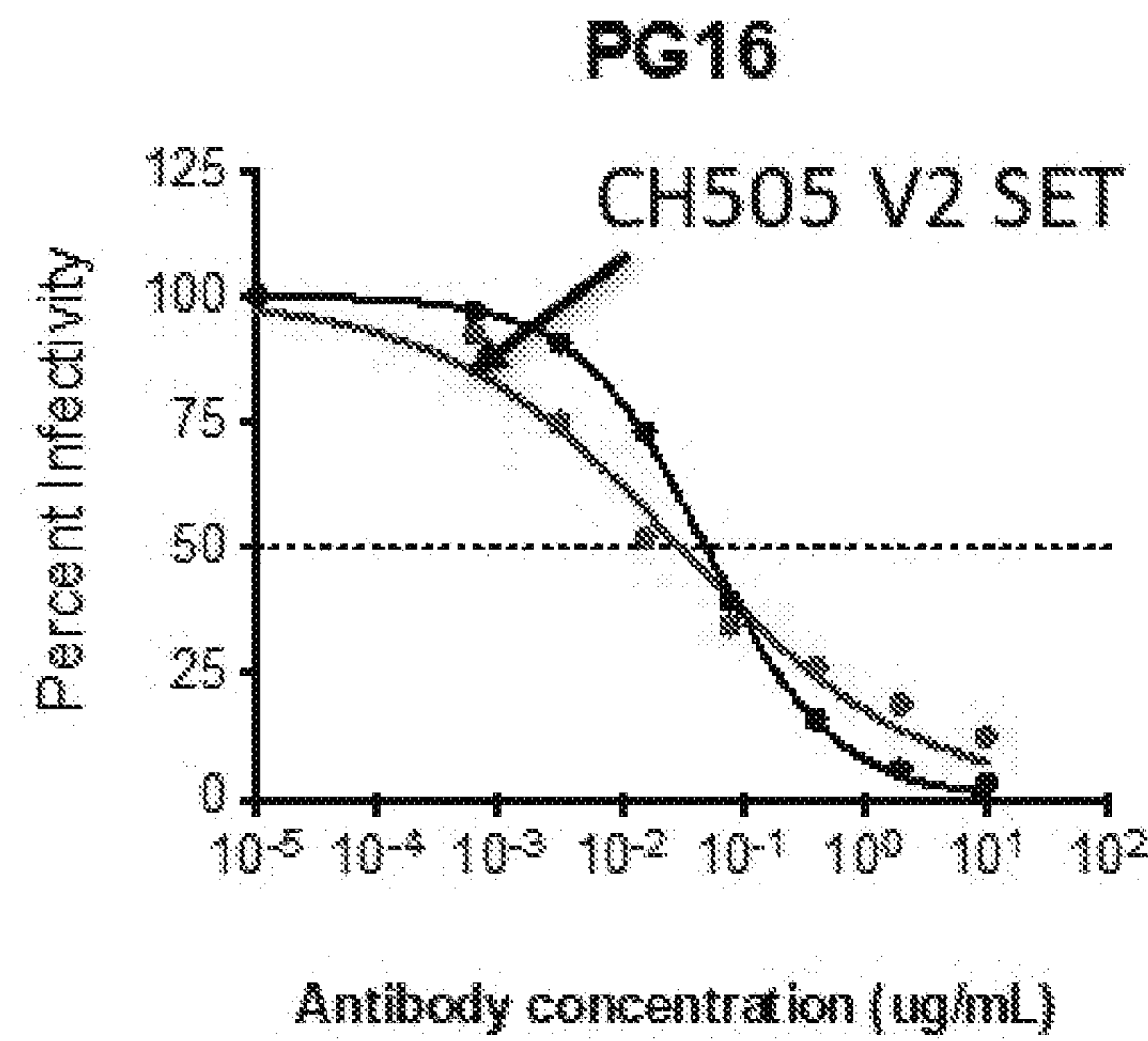


Figure 9D

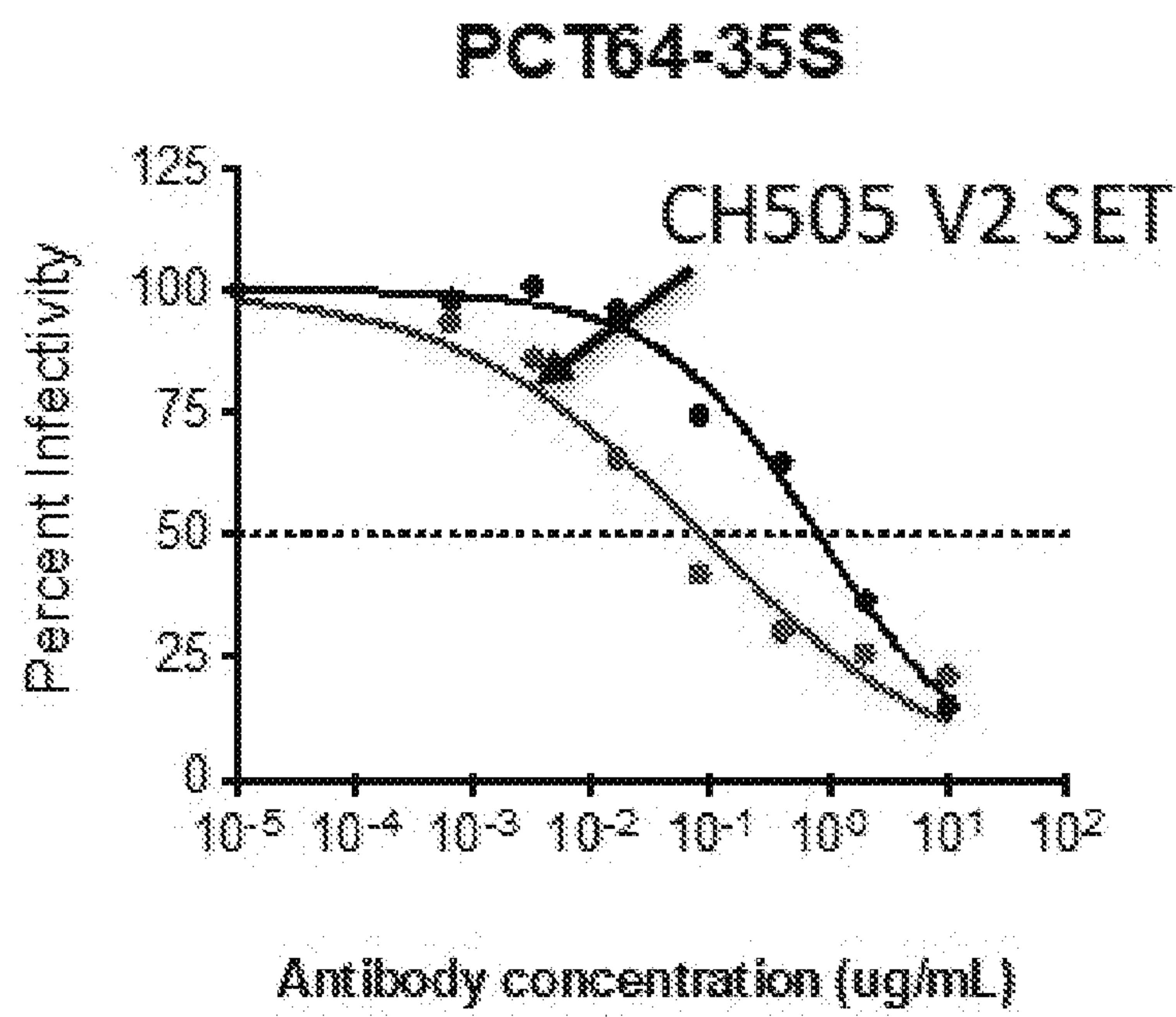


Figure 9E

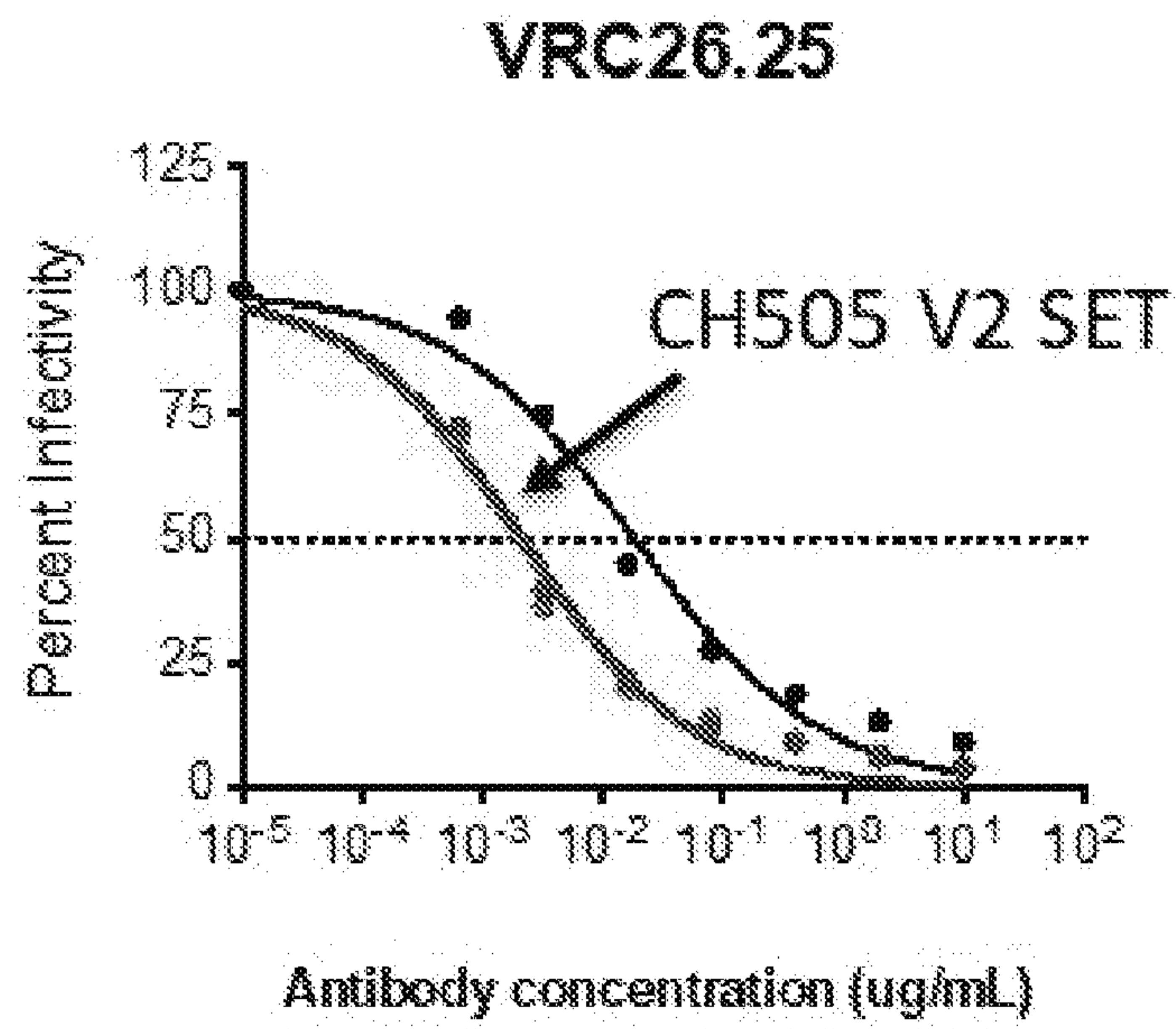


Figure 10A

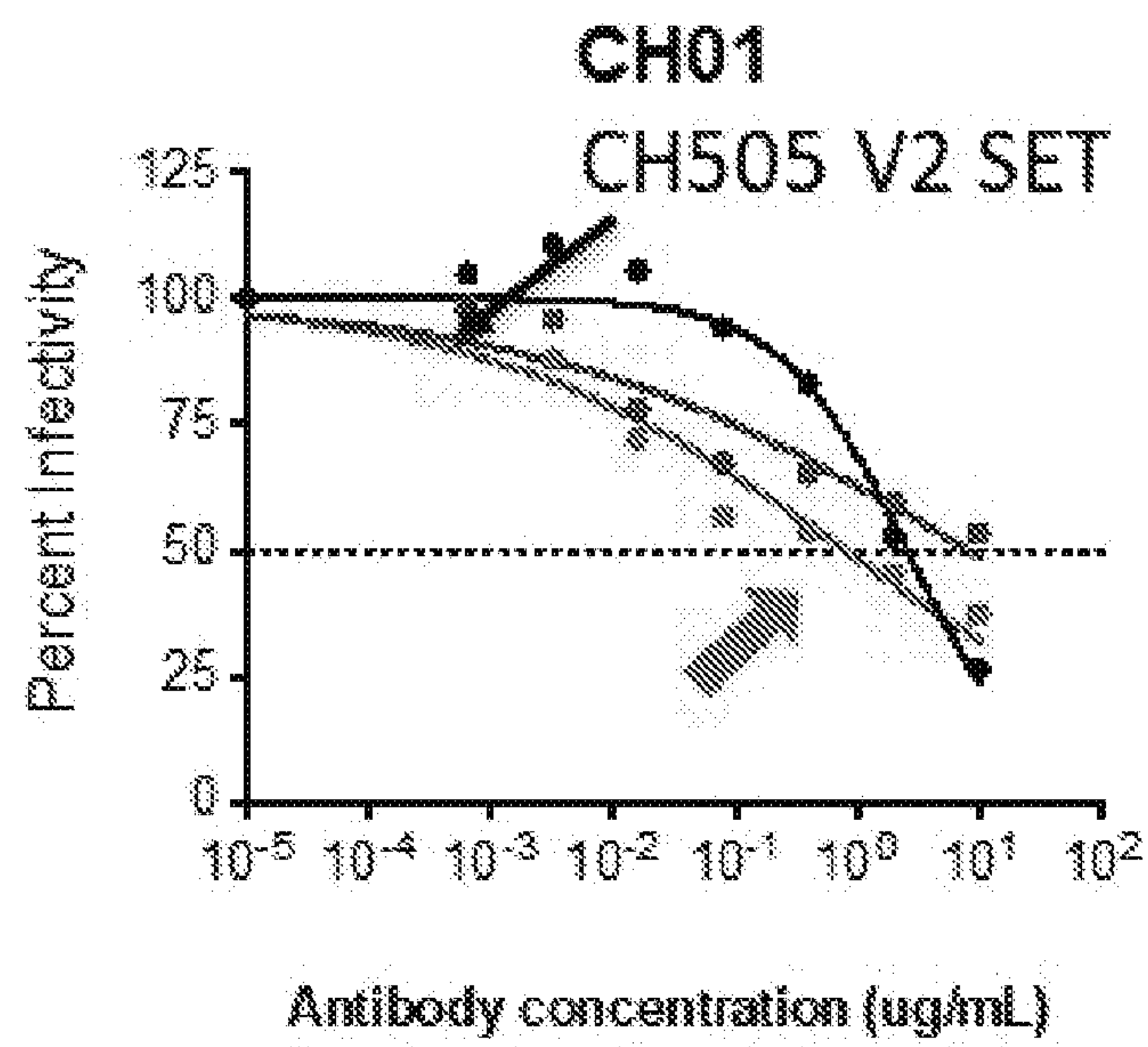


Figure 10B

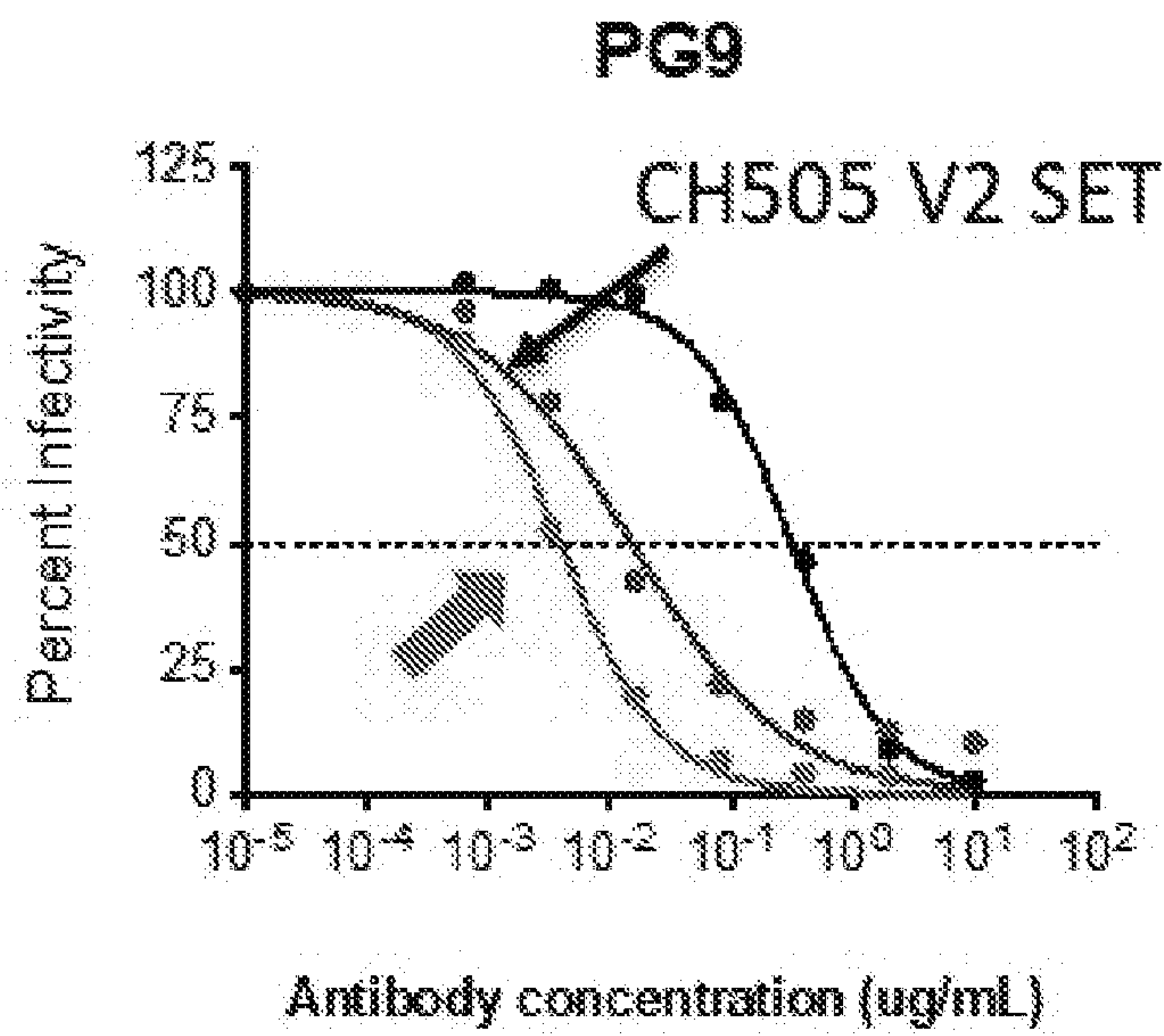


Figure 10C

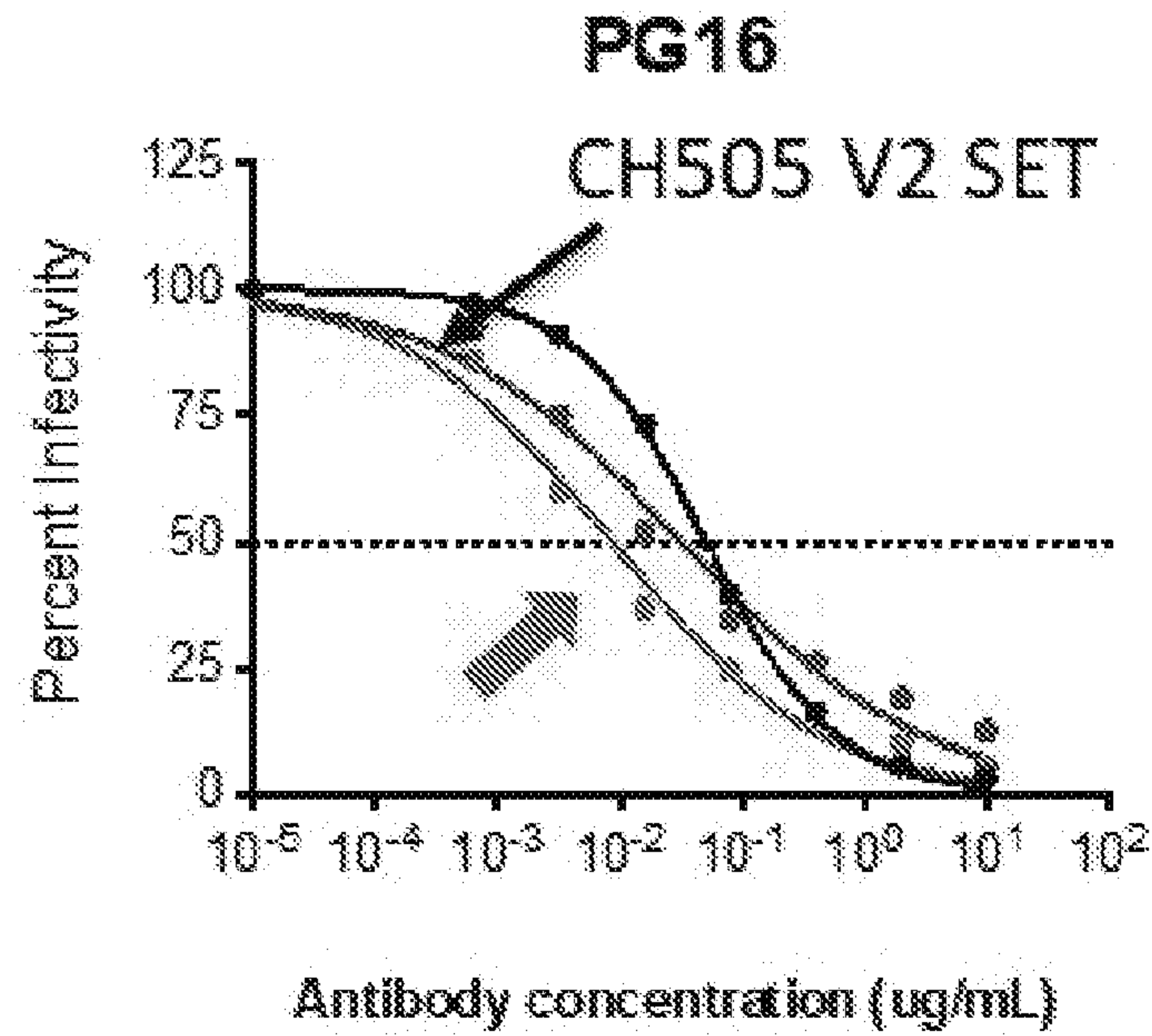


Figure 10D

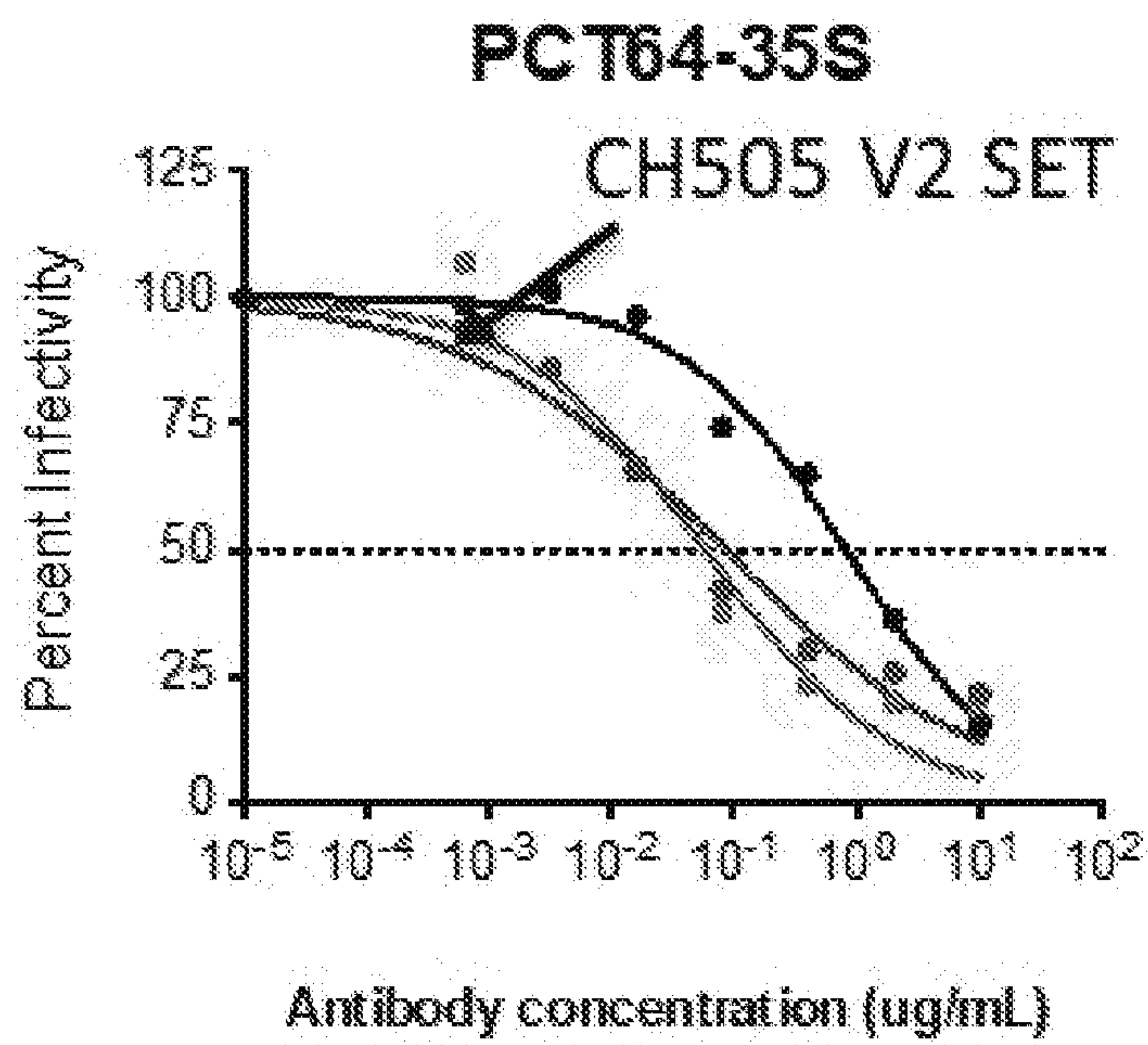


Figure 10E

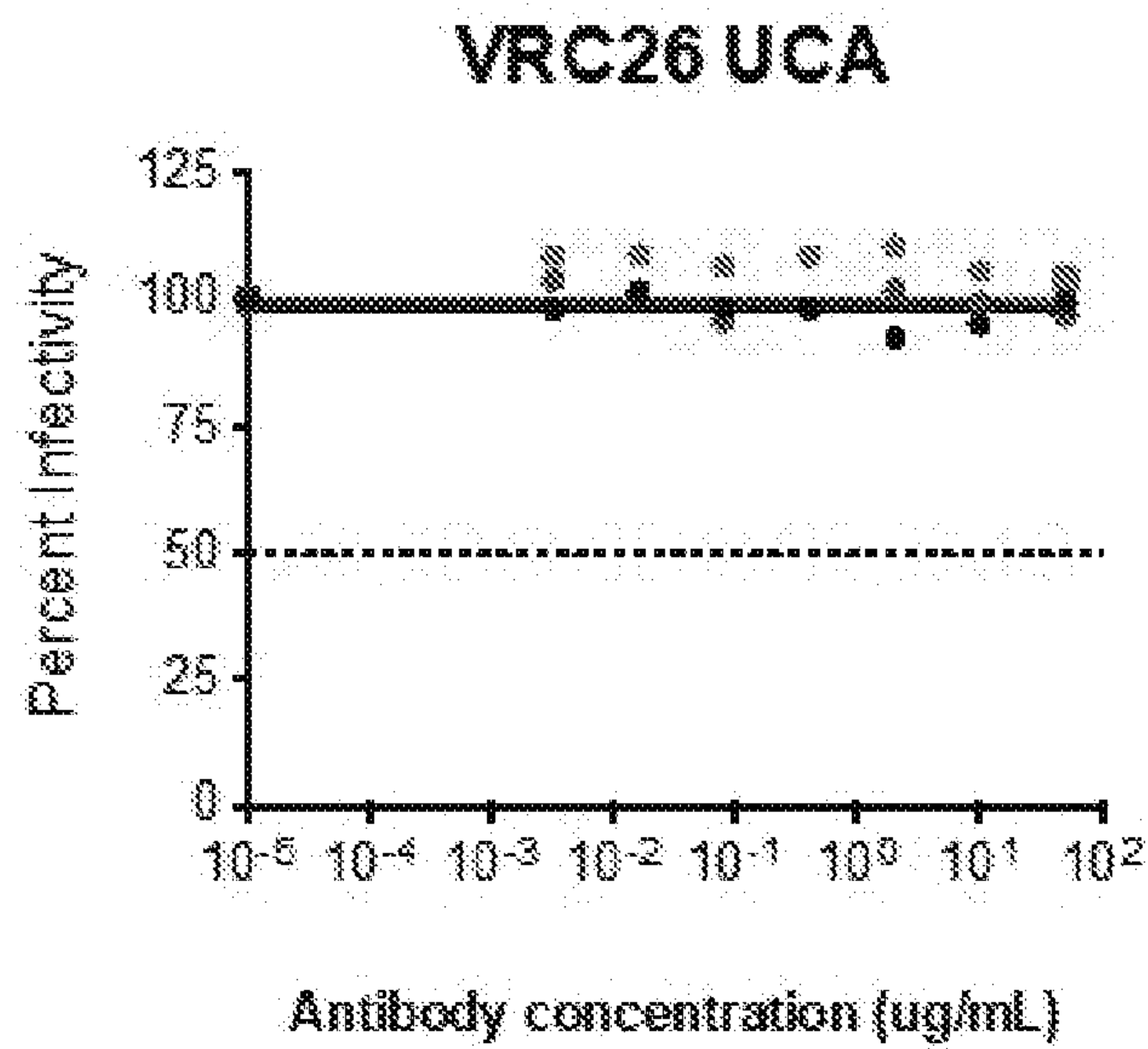


Figure 11A

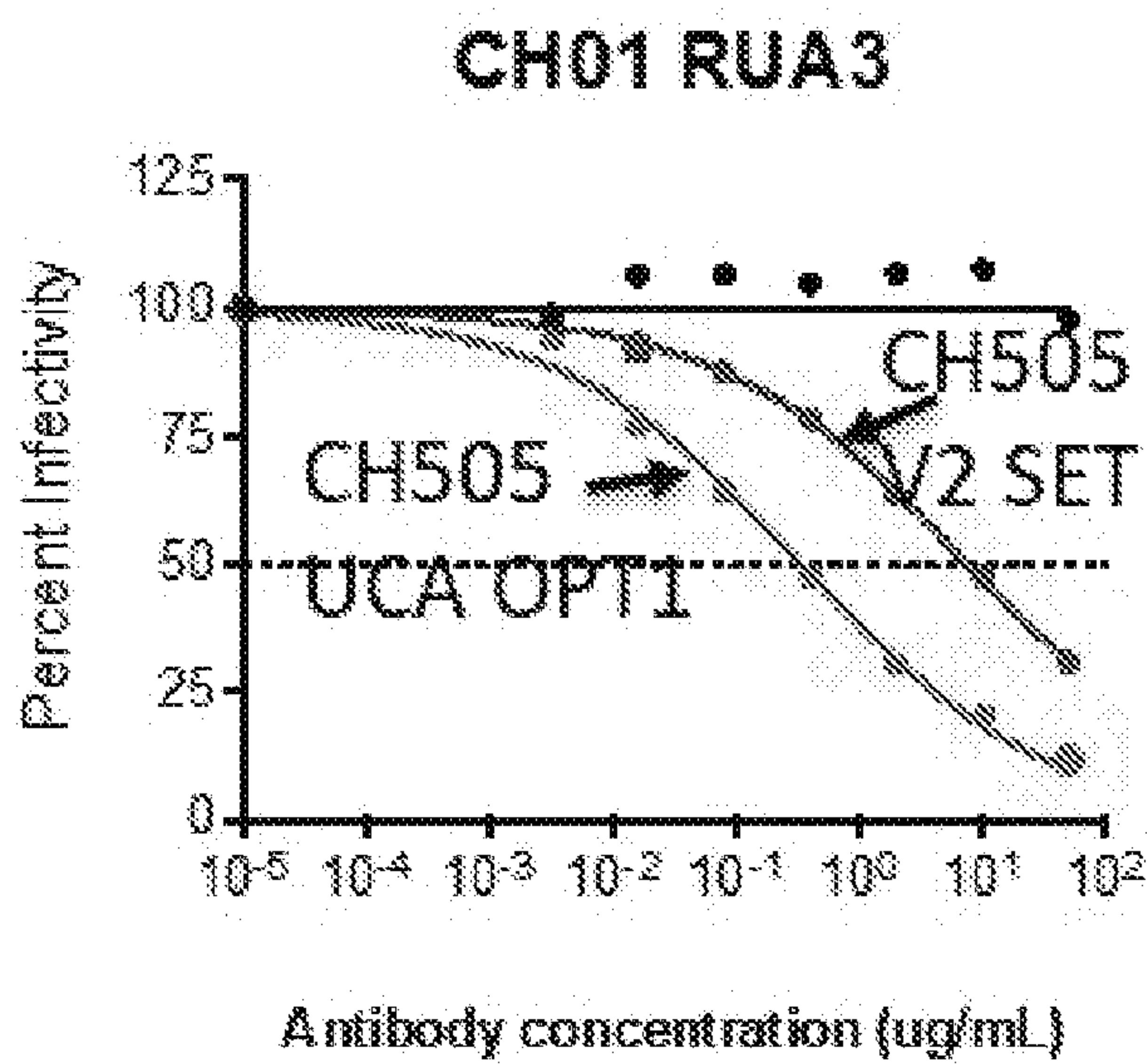


Figure 11B

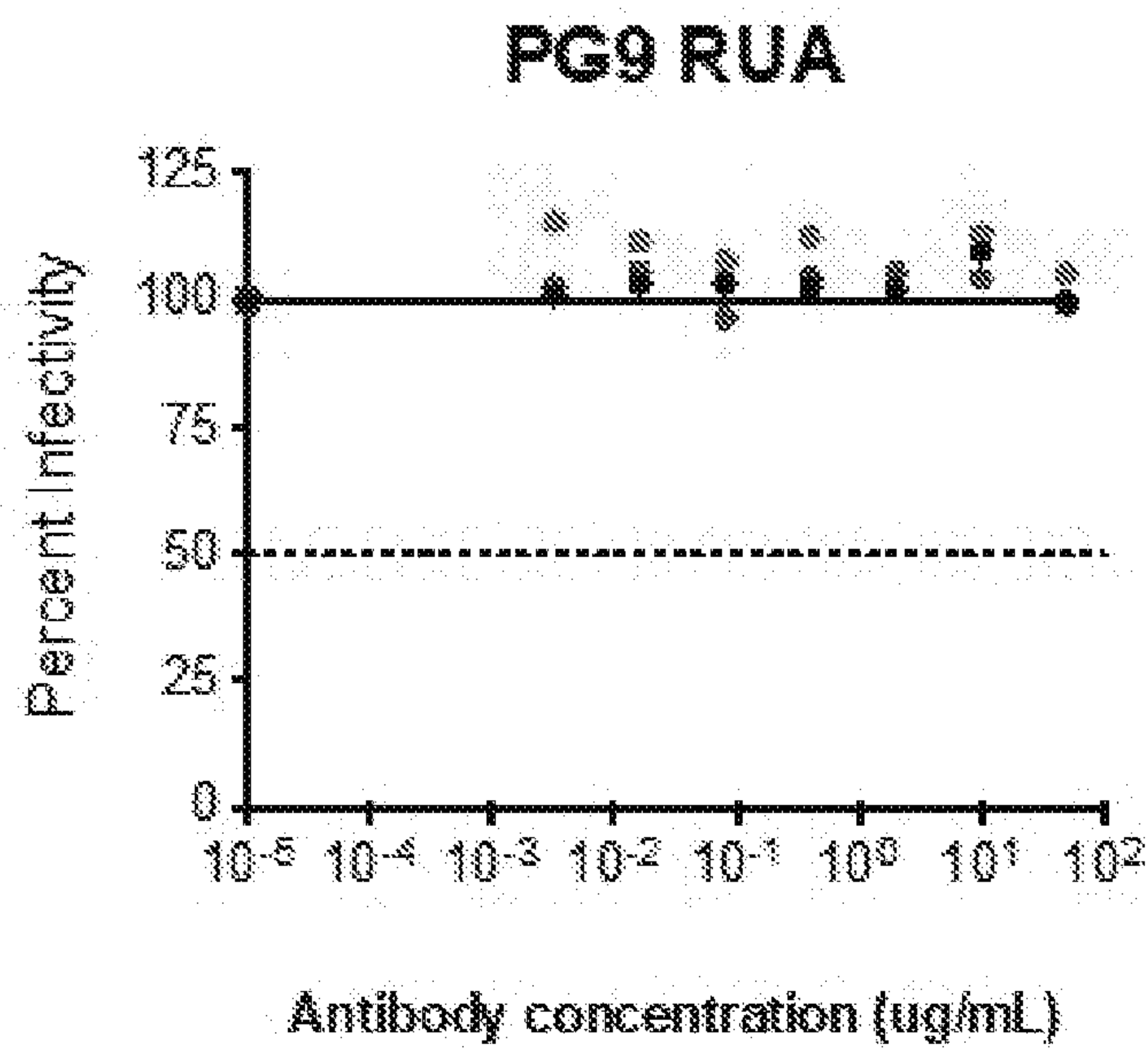


Figure 11C

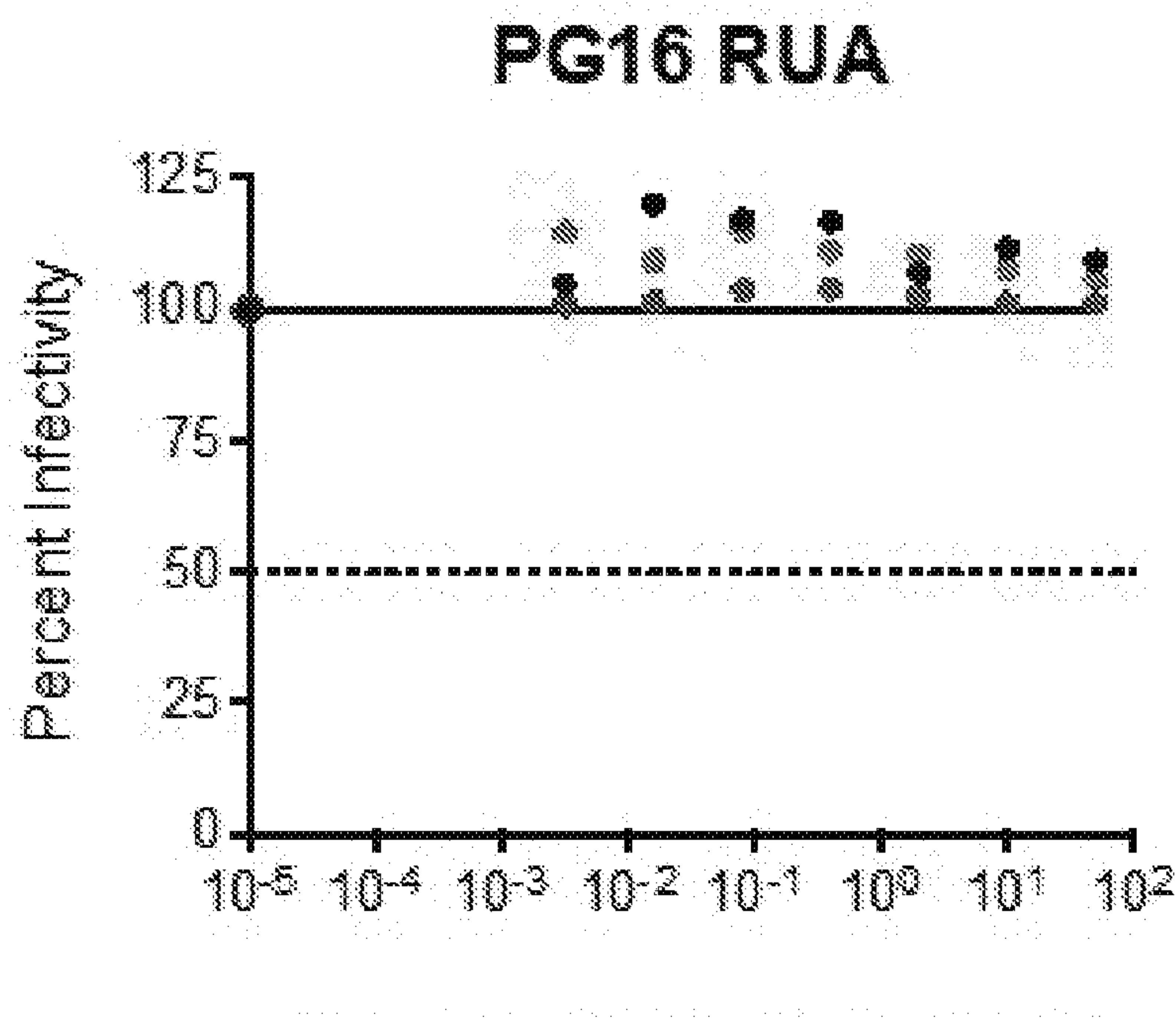


Figure 11D

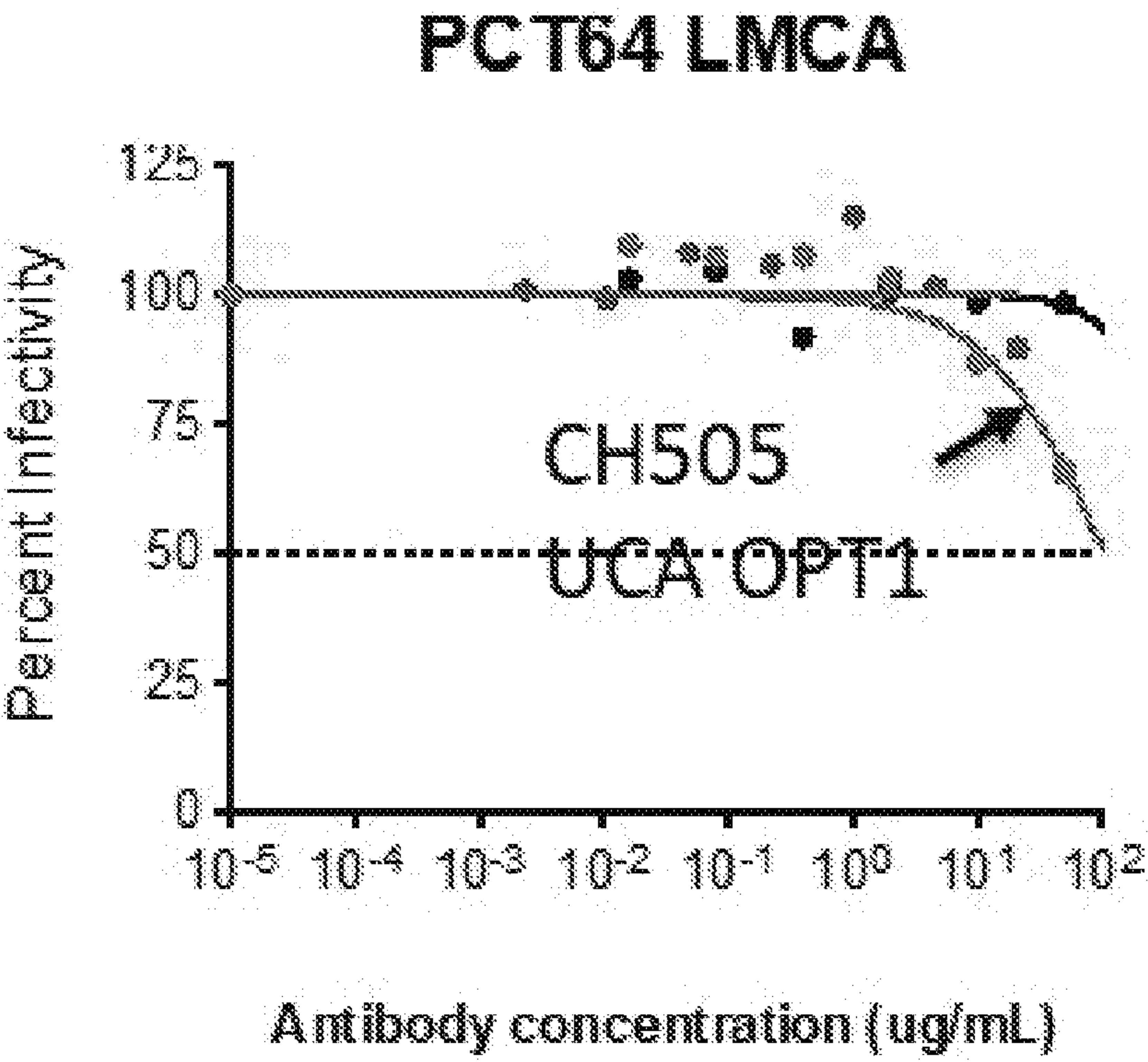


Figure 11E

Ab	WT	SET OPT	UCA OPT
CAP256.25	1.36	0.002	0.002
CH01	9.25	7.82	0.480
PG9	1.15	0.009	0.004
PG16	0.369	0.037	0.007
PCT64-35S	8.06	0.086	0.030
VRC26.25 UCA	>250	>50	>250
CH01 RUA3	>50	3.19	0.170
PG9 RUA	>50	>50	>50
PG16 RUA	>50	>50	>50
PCT64 LMCA	>250	>50	105

Figure 12

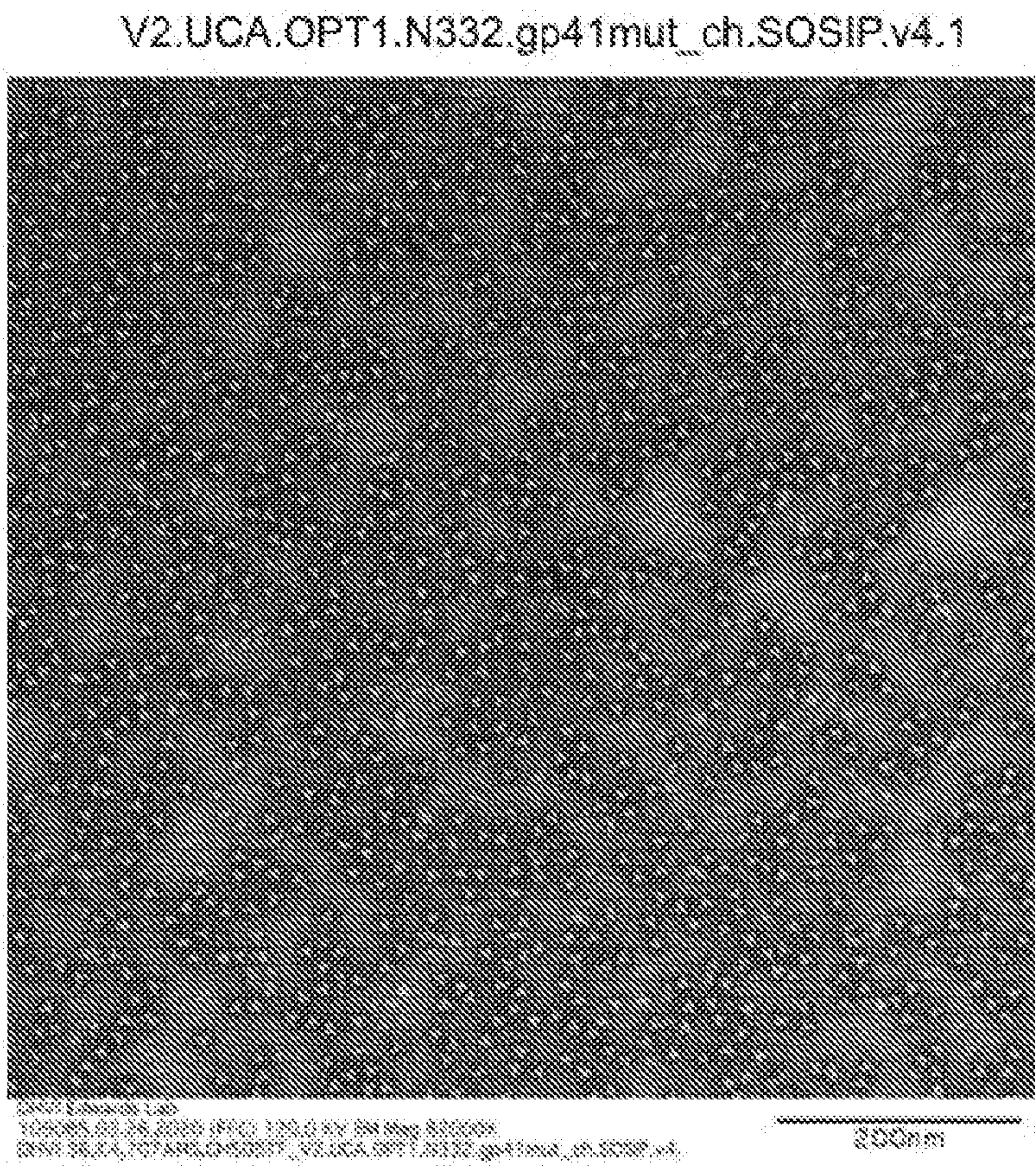


Figure 13A

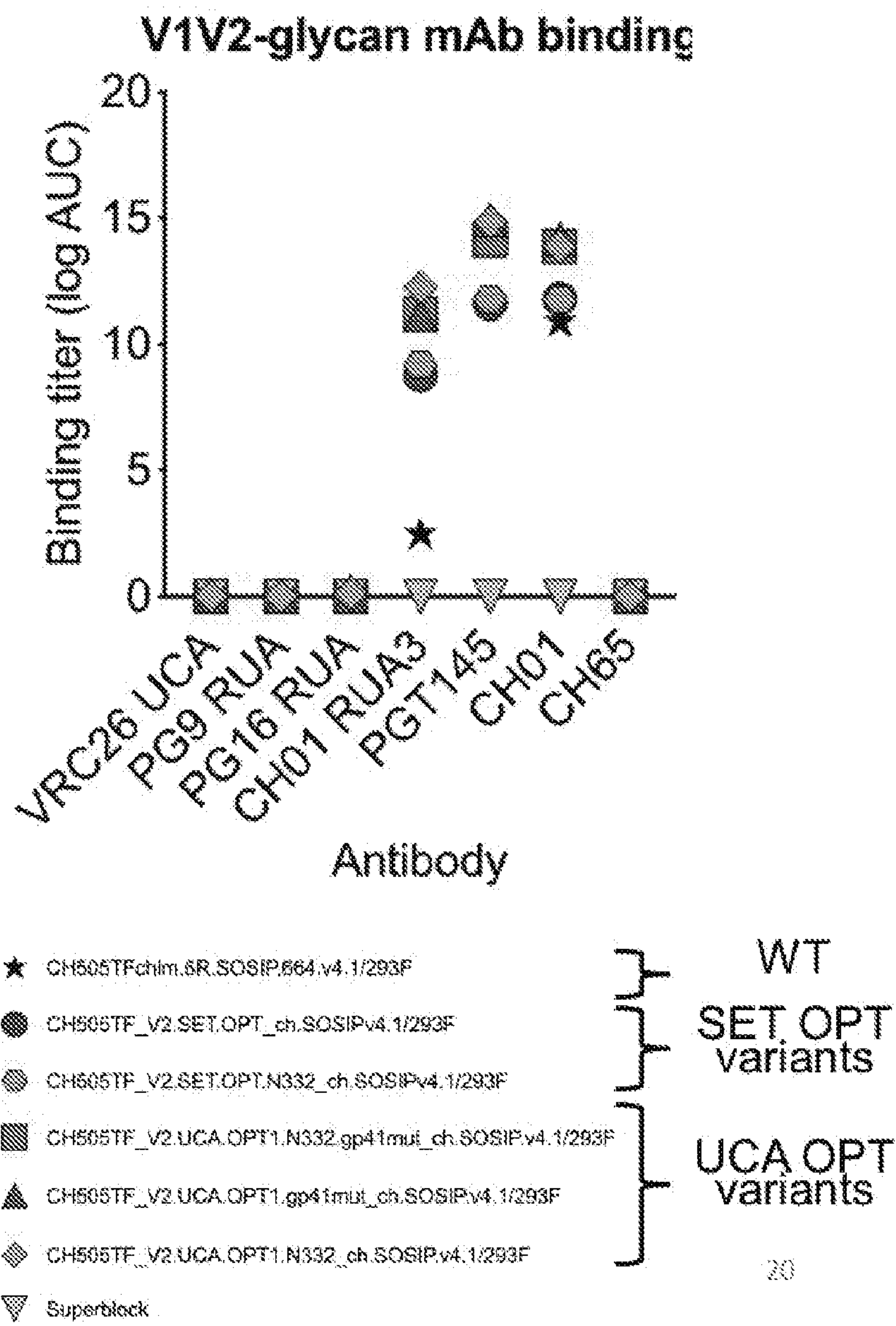


Figure 13B

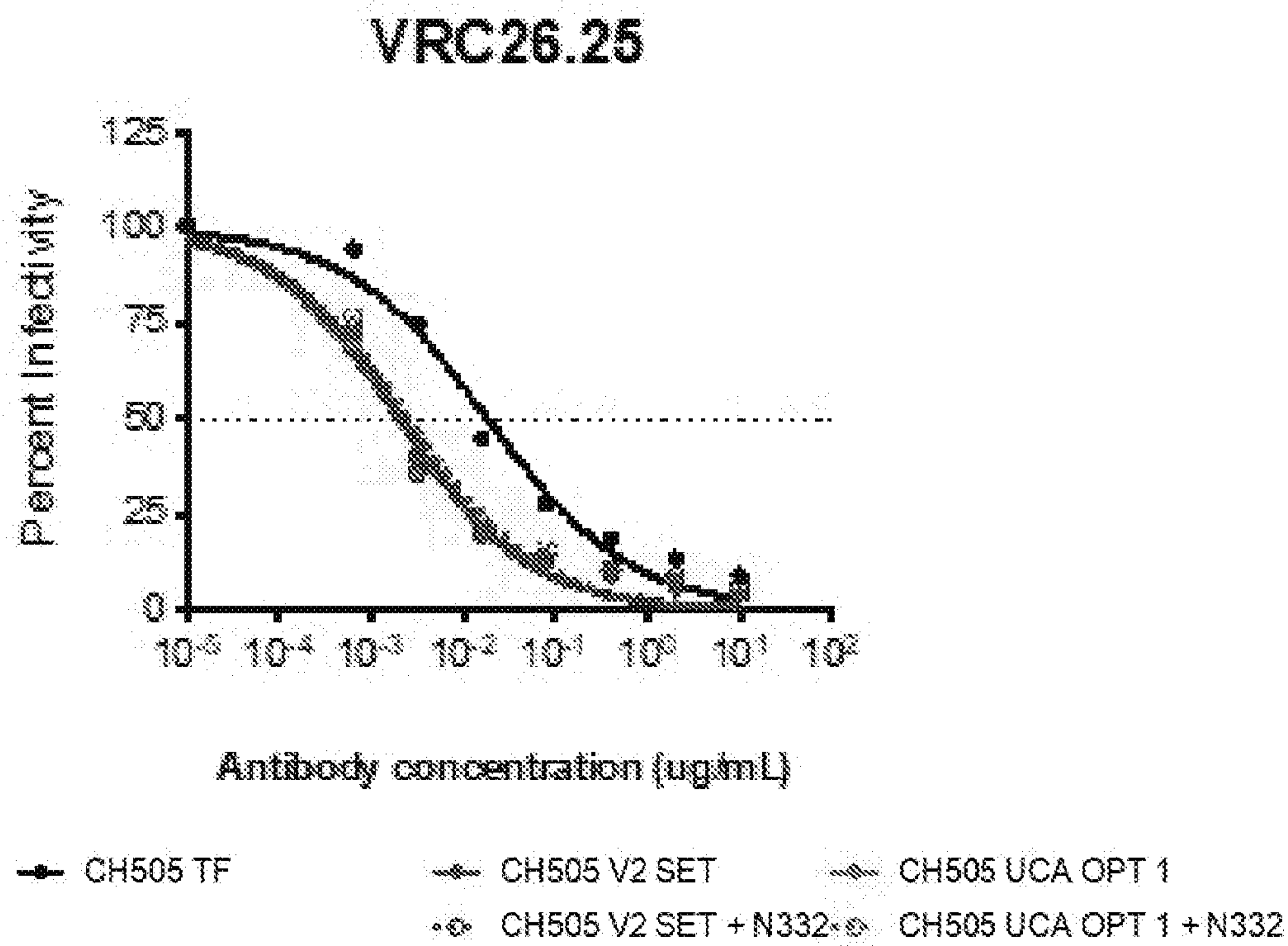


Figure 14A

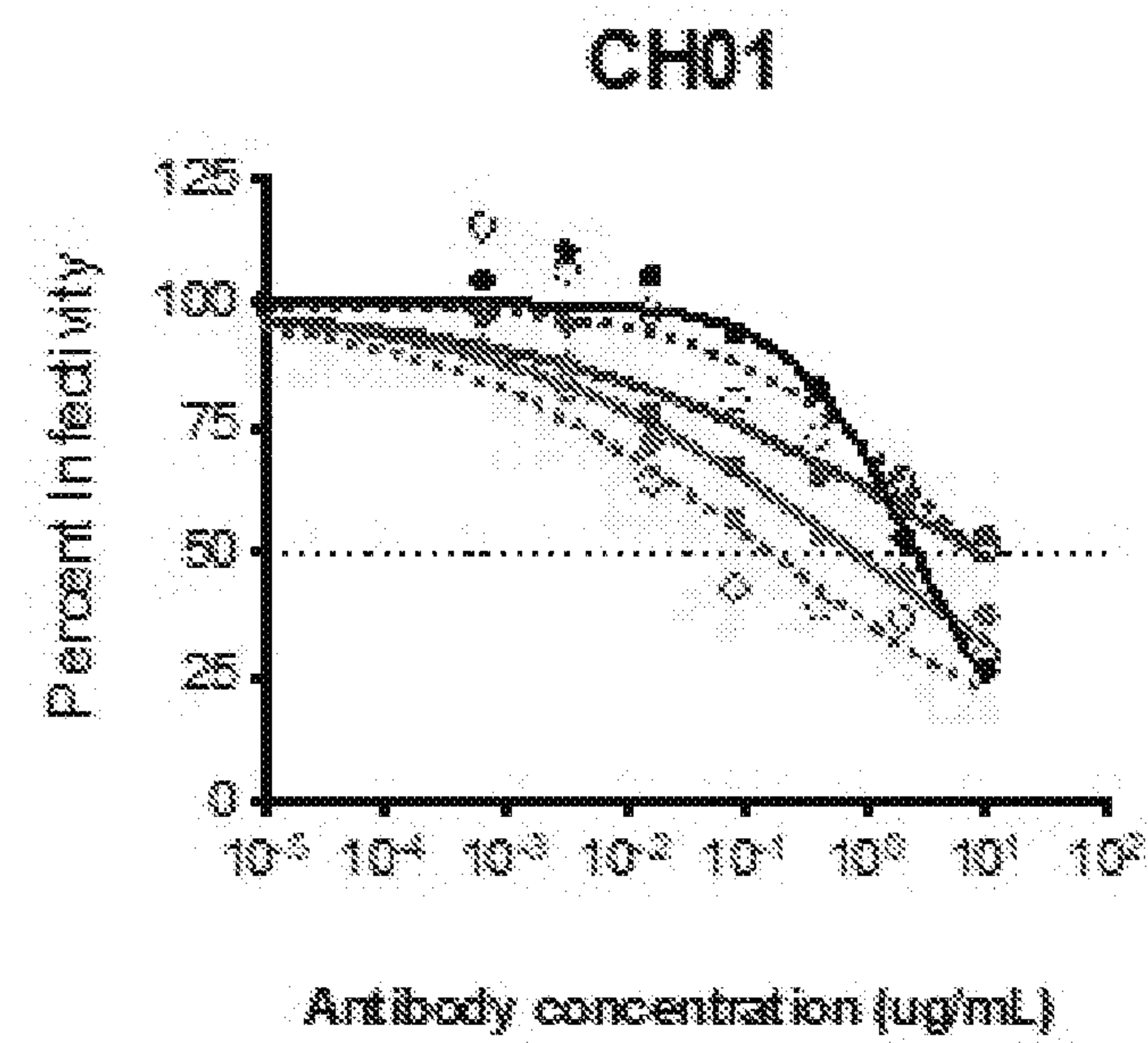


Figure 14B

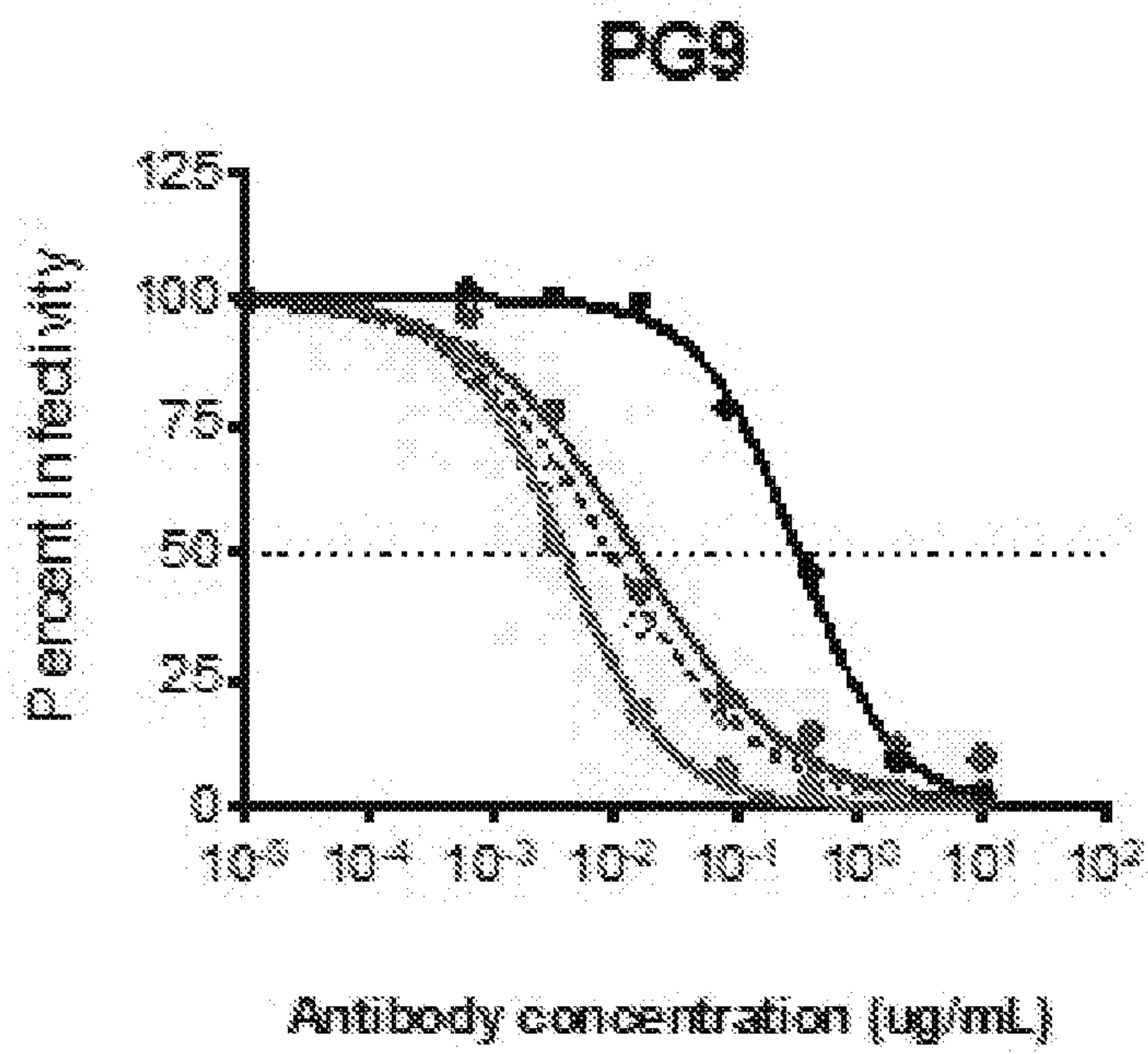


Figure 14C

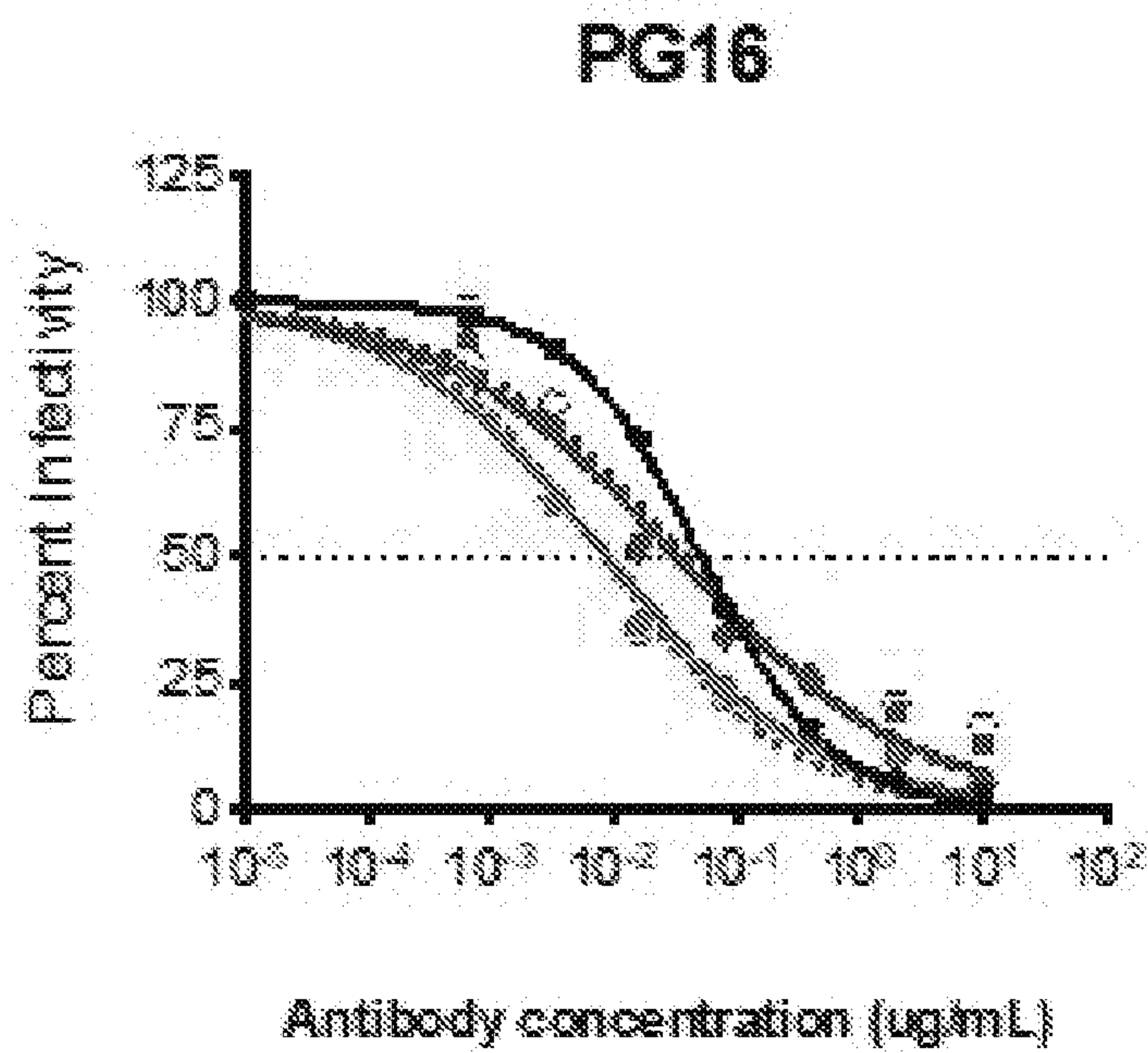


Figure 14D

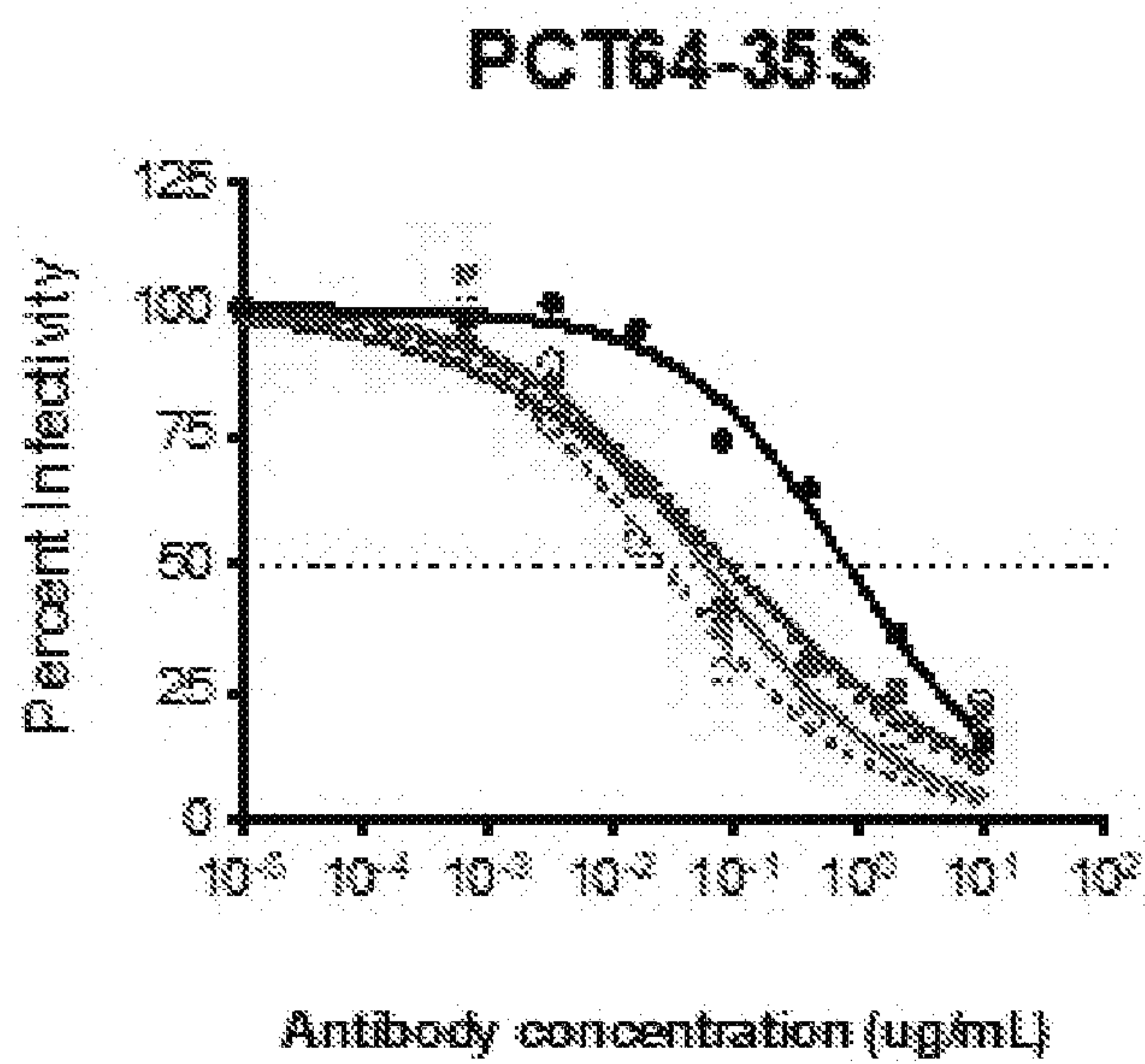


Figure 14E

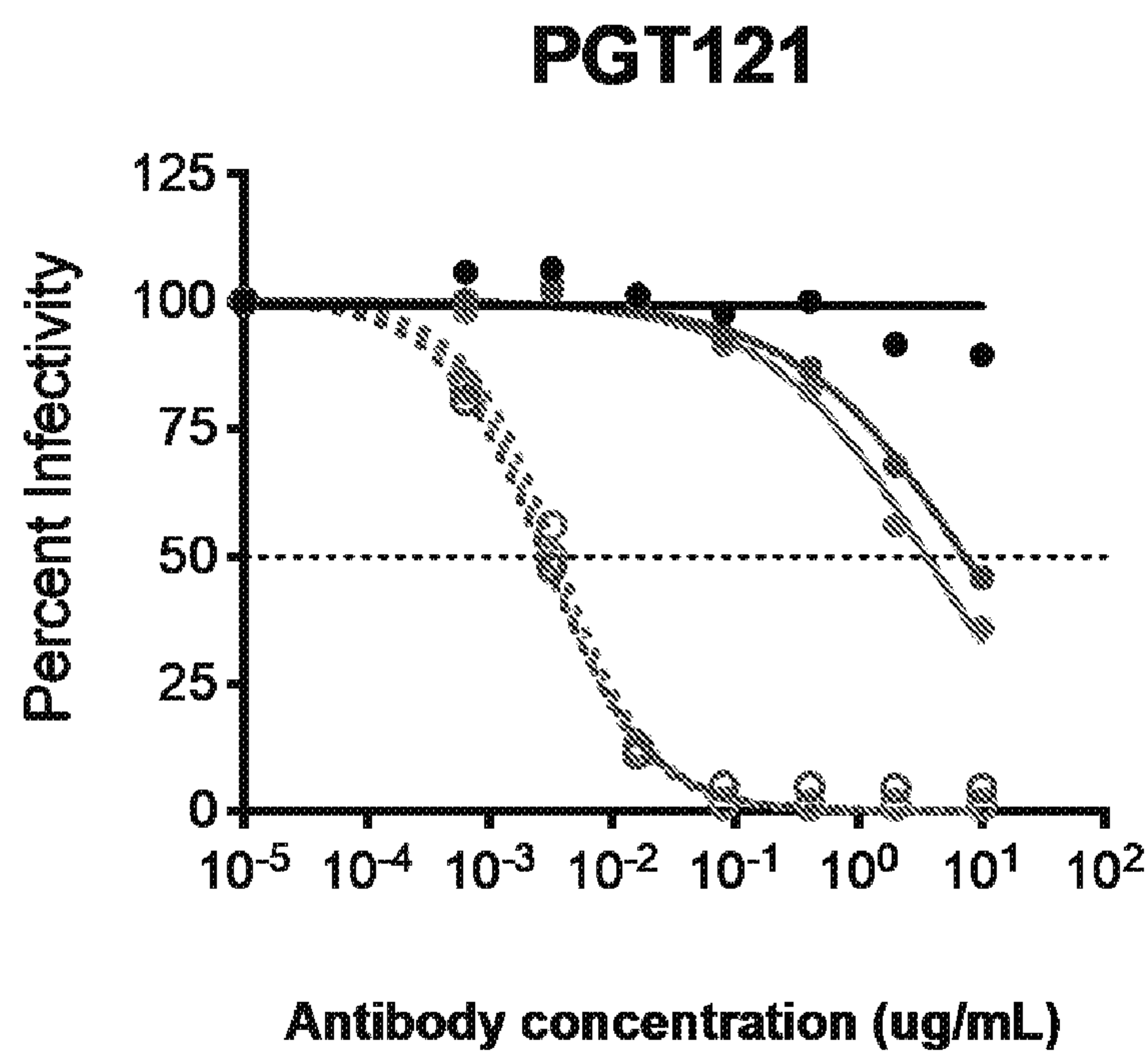


Figure 14F

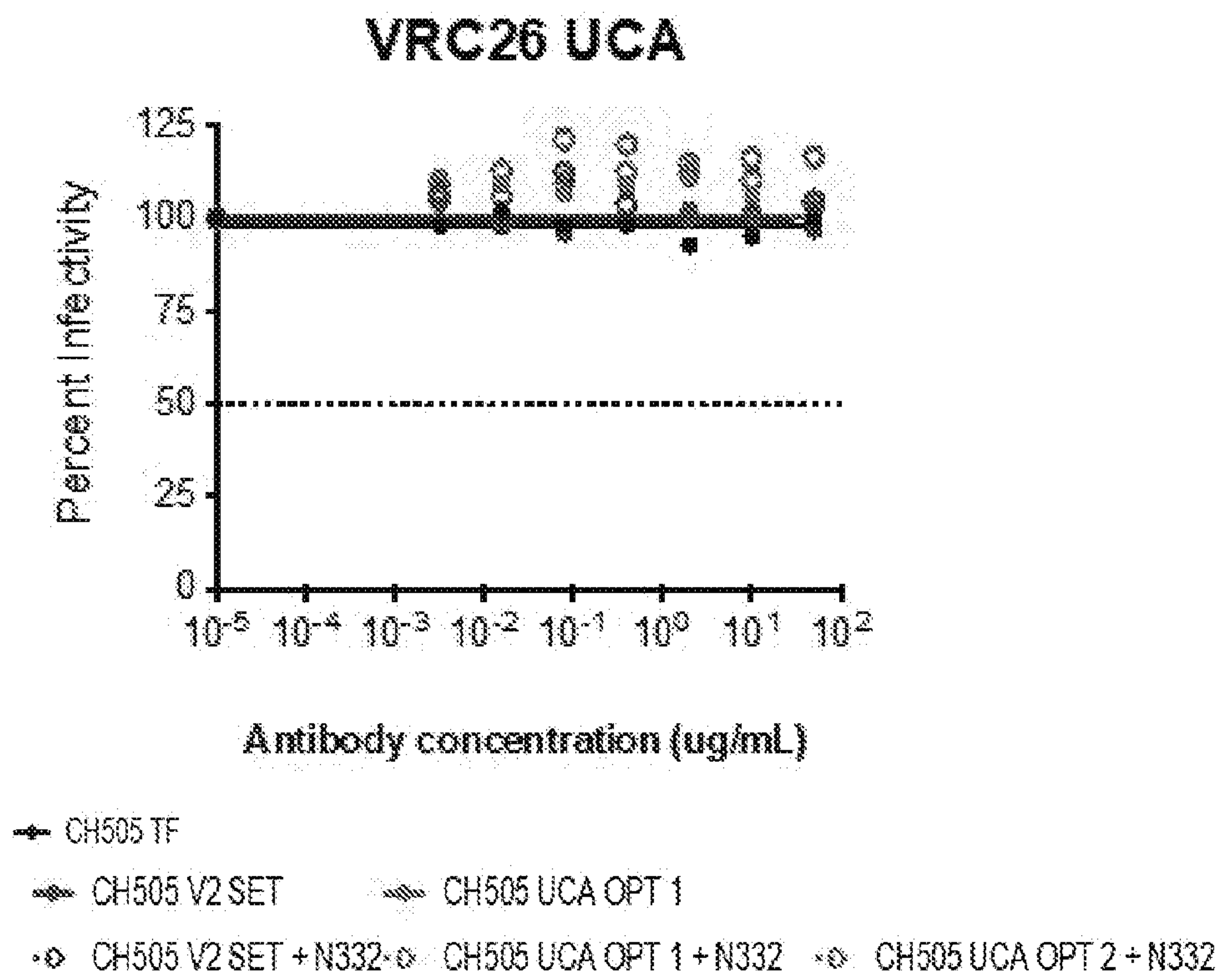


Figure 15A

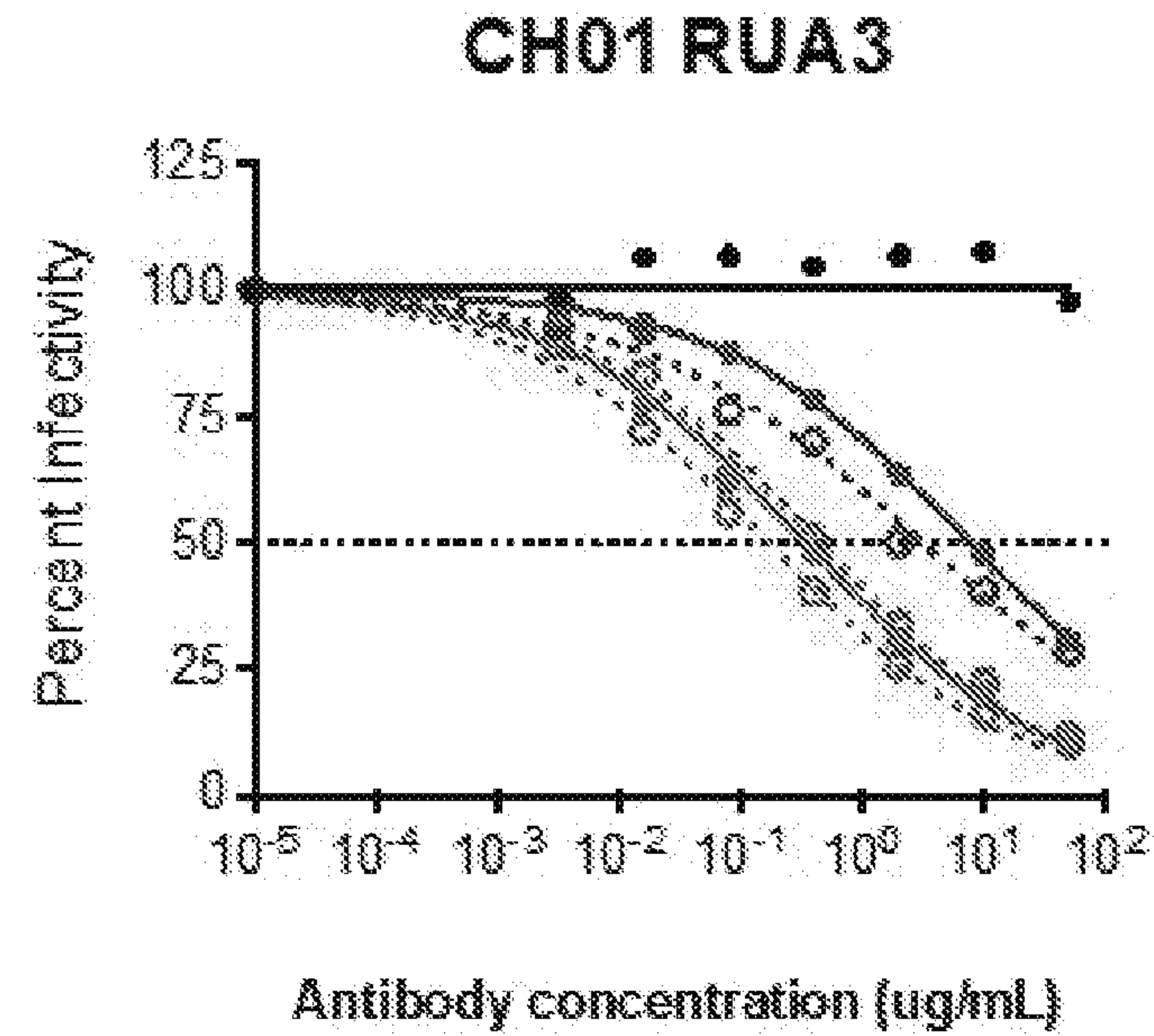


Figure 15B

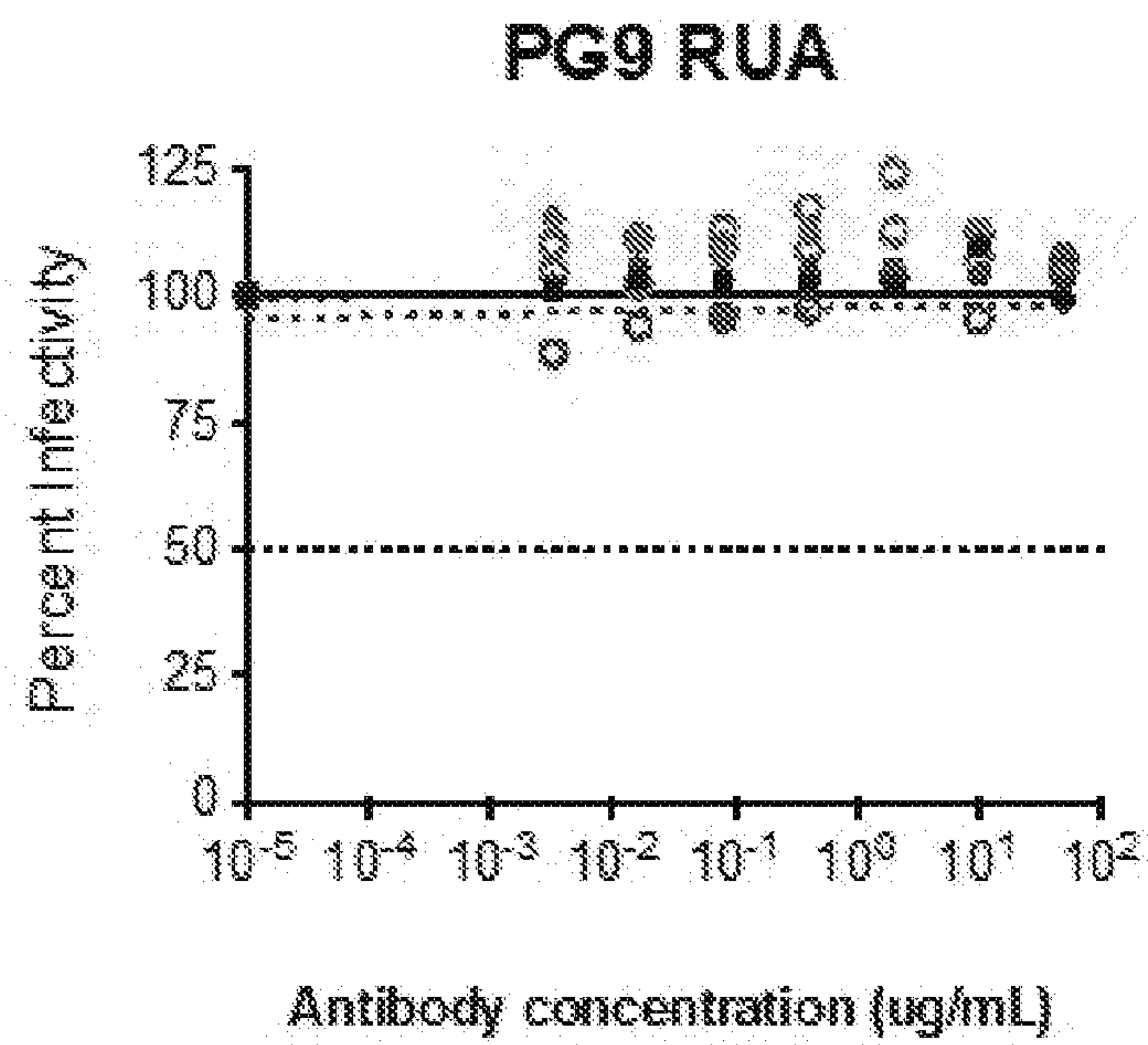


Figure 15C

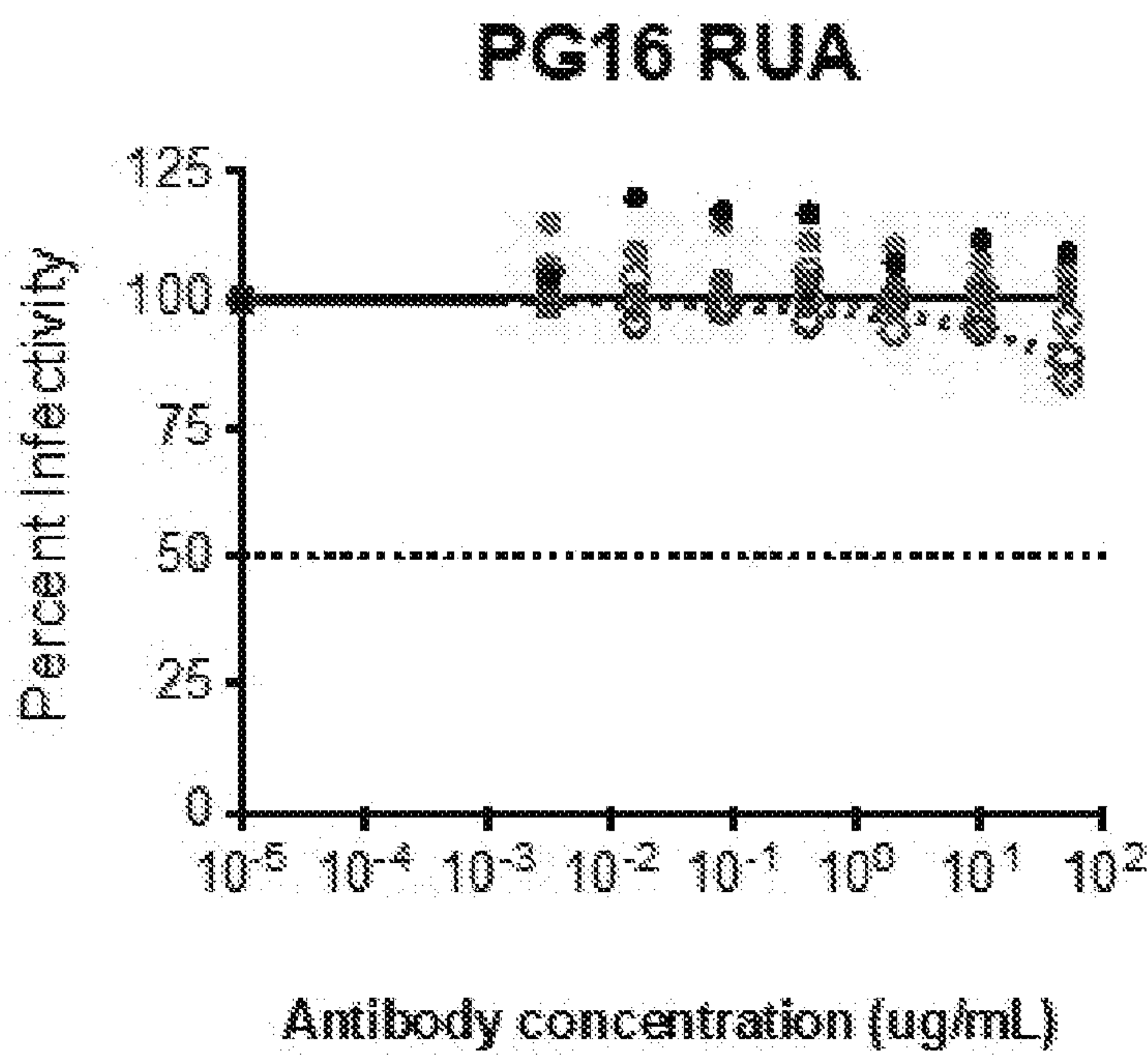


Figure 15D

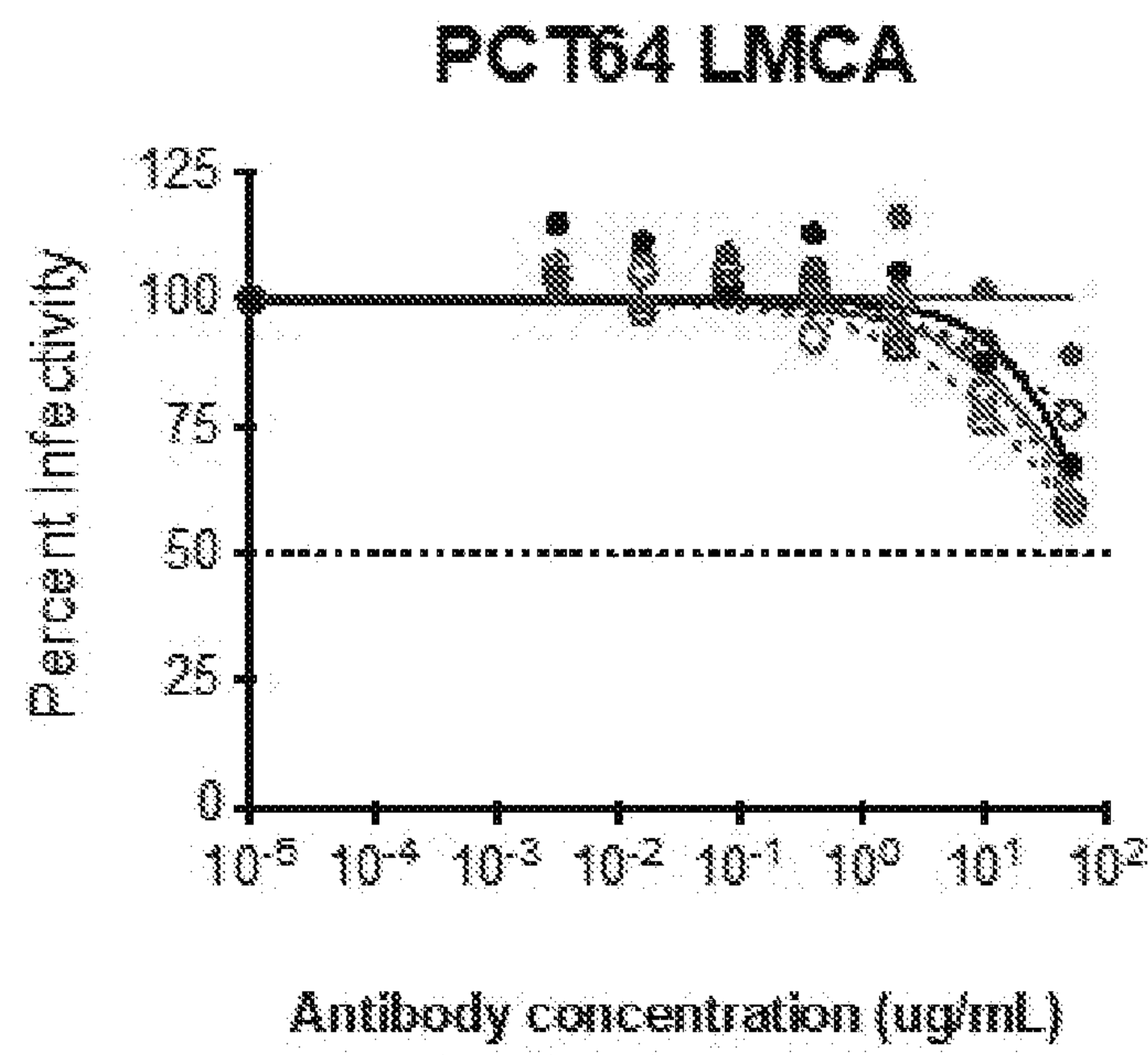


Figure 15E