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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
26 March 2020 (26.03.2020)

WIPO | PCT

(10) International Publication Number
WO 2020/061564 A1

(51) International Patent Classification:
A61K 39/21 (2006.01) *C07K 14/195* (2006.01)
A61K 39/385 (2006.01) *C07K 14/47* (2006.01)
B82Y 5/00 (2011.01) *C12N 9/10* (2006.01)
C07K 14/08 (2006.01) *C07K 14/02* (2006.01)

(21) International Application Number:
PCT/US2019/052419

(22) International Filing Date:
23 September 2019 (23.09.2019)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/735,188 23 September 2018 (23.09.2018) US

(71) Applicant: **THE UNITED STATES OF AMERICA, AS REPRESENTED BY THE SECRETARY, DEPARTMENT OF HEALTH AND HUMAN SERVICES** [US/US]; National Institutes of Health, Office of Technology Transfer, 6011 Executive Boulevard, Suite 325, MSC 7660, Bethesda, Maryland 20892-7660 (US).

(72) Inventors: **TSYBOVSKY, Yaroslav**; Frederick National Laboratory for Cancer Research, Leidos Biomedical Research, Inc., Post Office Box B, Frederick, Maryland 21702 (US). **XU, Kai**; NIH/NIAID, 40 Convent Dr., Building 40, Bethesd, Maryland 20892 (US). **MASCOLA, John**; NIH/NIAID, 40 Convent Dr., Building 40, Bethesda, Maryland 20892 (US). **KWONG, Peter**; NIH/NIAID, 40 Convent Dr., Building 40, Bethesda, Maryland 20892 (US). **OU, Li**; NIH/NIAID, 40 Convent Dr., Building 40, Bethesda, Maryland 20892 (US). **ZHANG, Baoshan**; NIH/NIAID, 40 Convent Dr., Building 40, Bethesda, Maryland 20892 (US). **KONG, Wing-Pui**; NIH/NIAID, 40 Convent Dr., Building 40, Bethesda, Maryland 20892 (US). **CHAO, Cara W.**; Building 40, Room 4510, 40 Convent Dr., Bethesda, Maryland 20814 (US).

(74) Agent: **SCOTT, Gregory K.**; Klarquist Sparkman, LLP, One World Trade Center, Suite 1600, 121 SW Salmon Street, Portland, Oregon 97204 (US).

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

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

Declarations under Rule 4.17:
— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*
— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

Published:
— *with international search report (Art. 21(3))*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*
— *with sequence listing part of description (Rule 5.2(a))*

(54) Title: HIV-1 ENV FUSION PEPTIDE NANOPARTICLE CARRIER CONJUGATES AND THEIR USE

(57) Abstract: Embodiments of immunogenic conjugates including the HIV-1 Env fusion peptide and methods of their use and production are disclosed. In several embodiments, the immunogenic conjugates can be used to generate an immune response to HIV-1 Env in a subject, for example, to treat or prevent an HIV-1 infection in the subject.

WO 2020/061564 A1

HIV-1 ENV FUSION PEPTIDE NANOPARTICLE CARRIER CONJUGATES AND THEIR USE

CROSS REFERENCE TO RELATED APPLICATION

This application claims priority to U.S. Provisional Application No. 62/735,188, filed September 23, 2018, which is incorporated by reference in its entirety.

FIELD

This disclosure relates to immunogenic conjugates including HIV-1 envelope (Env) fusion peptides conjugated to a self-assembling protein nanoparticle carrier and their use to induce an immune response in a subject.

BACKGROUND

Millions of people are infected with HIV-1 worldwide, and 2.5 to 3 million new infections have been estimated to occur yearly. Although effective antiretroviral therapies are available, millions succumb to AIDS every year, especially in sub-Saharan Africa, underscoring the need to develop measures to prevent the spread of this disease.

An enveloped virus, HIV-1 hides from humoral recognition behind a wide array of protective mechanisms. The major envelope protein of HIV-1 is a glycoprotein of approximately 160 kD (gp160). During infection, proteases of the host cell cleave gp160 into gp120 and gp41. Gp41 is an integral membrane protein, while gp120 protrudes from the mature virus. Together gp120 and gp41 make up the HIV-1 Env spike, which is a target for neutralizing antibodies.

It is believed that immunization with an effective immunogen including epitopes of the HIV-1 Env glycoprotein can elicit a neutralizing response, which may be protective against HIV-1 infection. However, despite extensive effort, a need remains for agents capable of such action.

SUMMARY

This disclosure provides novel immunogenic conjugates for eliciting an immune response to HIV-1 Env in a subject.

The immunogenic conjugates comprise a self-assembling protein-nanoparticle carrier conjugated to HIV-1 Env fusion peptides. The self-assembling protein-nanoparticle carrier is comprised of a multimer of fusion proteins. Each fusion protein in the multimer comprises a self-assembling protein nanoparticle subunit fused to a heterologous carrier protein. The fusion proteins self-assemble to form the self-assembling protein-nanoparticle carrier. The HIV-1 Env fusion peptides conjugated to the self-assembling protein-nanoparticle carrier, comprise, from the N-terminus, the amino acid sequence of residue 512 to one of residues 514-521 of a human immunodeficiency virus type 1 (HIV-1) Envelope (Env) protein (according to the HXB2 numbering system). In some embodiments, the fusion proteins in the self-assembling protein

nanoparticle carrier further comprise a heterologous T-cell helper epitope. The immunogenic conjugate can be used elicit an immune response to HIV-1 Env in a subject.

Immunogenic compositions including a disclosed immunogenic conjugate are also provided. The composition may be a pharmaceutical composition suitable for administration to a subject, and may also be contained in a unit dosage form. The composition can further include an adjuvant.

Methods of generating an immune response to HIV-1 Env protein in a subject are disclosed, as are methods of treating, inhibiting or preventing an HIV-1 infection in a subject. In such methods a subject, such as a human subject, is administered an effective amount of a disclosed immunogenic conjugate to elicit the immune response. In several embodiments, the method comprises a prime-boost immunization protocol, where a disclosed immunogenic conjugate is used for the prime immunization. The subject can be, for example, a human subject at risk of or having an HIV-1 infection.

The foregoing and other features and advantages of this disclosure will become more apparent from the following detailed description of several embodiments which proceeds with reference to the accompanying figures.

BRIEF DESCRIPTION OF THE FIGURES

FIGs. 1A-1I depict embodiments of the self-assembling protein nanoparticle carrier disclosed herein conjugated (FIGs. 1C-1I) or not (FIGs. 1A and 1B) to HIV-1 Env fusion peptides. As shown in FIG. 1A, the self-assembling protein nanoparticle carrier is a multimer of fusion proteins, each including a self-assembling protein nanoparticle subunit fused to a heterologous carrier protein. In some embodiments, the fusion protein can further include a T-cell helper epitope (FIG. 1B), which is then included in the self-assembling protein nanoparticle carrier. The location of the T-cell-helper epitope can be varied in the fusion protein. FIGs. 1C-1I, the HIV-1 Env fusion peptides (FP) are conjugated to the self-assembling protein nanoparticle carrier. The HIV-1 Env fusion peptides can be conjugated to any suitable aspect of the self-assembling protein nanoparticle carrier. In some instances, sulfosuccinimidyl (4-iodoacetyl)aminobenzoate (Sulfo-SIAB) conjugation chemistry is used to conjugate the HIV-1 Env fusion peptides to exposed lysine residues of the self-assembling protein nanoparticle carrier. FIGs. 1G-1I illustrate additional embodiments that further include a targeting moiety that targets the immune system in a subject to enhance the immune response to the HIV-1 Env fusion peptide on the immunogenic conjugate. The depictions in FIGs. 1A-1I are for illustration purposes and are not drawn to scale and do not necessarily show the number or relative location of self-assembling protein nanoparticle subunits, carrier proteins, HIV-1 Env fusion peptides, and T-cell helper epitopes that are present in a disclosed immunogenic conjugate.

FIG. 2 shows a set of images illustrating structural differences between KLH nanoparticles and KLH subunits, and a graph presenting data showing that immunization with KLH nanoparticles conjugated to FP8 peptide (AVGIGAVF, residues 1-8 of SEQ ID NO: 1) elicits a much greater immune response to the HIV-1 Env trimer than immunization with KLH subunit conjugated to FP8 peptide.

FIGs. 3A-3C shows a nanoparticle carrier assembly through genetic fusion of LS nanoparticle subunit and rTT carrier. FIG. 3A. Schematic of the fusion protein used to produce genetically fused rTT-LS nanoparticle. FIG. 3B. SEC profile of purified rTT-LS nanoparticle. FIG. 3C. Electron micrographs of genetically fused rTT-LS nanoparticle carrier shows particle species.

FIG. 4 shows electron micrographs of genetically fused rTT-LS nanoparticle carrier with a IgG hinge linking the rTT and LS subunit.

FIG. 5 shows electron micrographs for another example of a genetically fused nanoparticle carrier, formed from subunits of H influenzae protein D fused to phosphopantetheine adenylyltransferase nanoparticle subunit fused to rTT (HiD-6CCQ-rTT). The sequence of the fusion protein used to generate these nanoparticle carrier is provided as SEQ ID NO: 179. The observed particles were generally consistent in size and shape with the known phosphopantetheine adenylyltransferase crystal structure (PDB 6CCQ).

FIGs. 6A-6C shows a nanoparticle carrier assembled through isopeptide bond fusion of lumazine synthase nanoparticle subunit and rTT carrier. FIG. 6A. Schematic of the lumazine synthase-spytag and rTT-spycatcher fusion proteins used to produce isopeptide bond-fused rTT-LS nanoparticle. Subsequent to formation of the rTT-LS nanoparticle, HIV-1 fusion peptide (FP8) was conjugated to the nanoparticle-carrier by a PEG linker. FIG. 6B. Coomassie stained SDS-PAGE shows the individual purified proteins. FIG. 6C. SEC profile of purified rTT-LS nanoparticle.

FIG. 7 is a series of electron micrograph images of the purified rTT-SpyC fusion protein, the LS-SpyT nanoparticle, the LS-SpyT nanoparticle joined to the rTT-SpyC fusion protein (LS-Spy-rTT), and the LS-SpyT nanoparticle joined to the rTT-SpyC fusion protein further conjugated to HIV-1 Env fusion peptide FP8v1 by a PEG linker (LS-Spy-rTT-FP8v1/PEG2).

FIG. 8 shows results of isothermal calorimetry assays to determine the number of HIV-1 Env fusion peptides conjugated to monomeric rTT (FP-rTT) compared to the number of HIV-1 Env fusion peptides conjugated to the LS-SpyT nanoparticle joined to the rTT-SpyC fusion protein (LS-Spy-rTT-FP8v1/PEG2 nanoparticle carrier). The results show that each FP-rTT monomer entity has six competent VRC34 Fab binding sites, whereas each LS-Spy-rTT-FP8v1/PEG2 nanoparticle carrier has 152 – 402 competent VRC34.01 Fab binding sites.

FIG. 9 depicts an immunization protocol used to assess the LS-Spy-rTT-FP8v1/PEG2 nanoparticle carrier. For the first three immunizations (weeks 0, 3, and 6), mice received a 25 μg dose of either FP8v1-rTT monomer (Groups 1 and 2) or LS-Spy-rTT-FP8v1/PEG2 nanoparticle carrier (Groups 3 and 4). For the following three immunizations, mice received a 25 μg dose of either BG505 DS-SOSIP trimer (Groups 1 and 3) or the BG505 DS-SOSIP trimer conjugated to a lumazine synthase nanoparticle (Groups 2 and 4). Adjuplex was used as adjuvant for each immunization. Blood was drawn at weeks 0, 2, 5, 8, 11, 14, and 17.

FIGs. 10A-10C show binding and neutralization characteristics for sera from FP-immunized mice. FIG. 10A. Week 2 and Week 5 sera was assessed for FP binding by octet binding assay. FIG. 10B. Week, 2, 5, and 8 sera was assessed for BG505 trimer binding by ELISA. FIG. 10C. Week 17 sera was assessed

for neutralization of BG505 virus with a mutation to remove glycan 611, as this viral variant is more sensitive to fusion peptide-directed antibodies (Kong *et al.* Science 352, 828-833, 2016).

FIG. 11 shows a SDS-PAGE gel illustrating purification of an encapsulin nanoparticle subunit fused to a spytag. The encapsulin subunit includes G53C-R94C mutations to introduce a disulfide bond that stabilizes nanoparticles formed for the subunit.

FIG. 12 shows a SDS-PAGE gel illustrating purification of an encapsulin nanoparticle subunit fused to a spytag (EN-spytag), rTT carrier fused to a spycatcher moiety (rTT-spyC), and the encapsulin-rTT fusion (rTT-EN) formed from these two molecules. The encapsulin subunit includes G53C-R94C mutations to introduce a disulfide bond that stabilizes nanoparticles formed for the subunit.

FIG. 13 shows a series of electron micrograph images of the purified encapsulin-spytag, rTT-spy-encapsulin fusion, and FP8v1-rTT-spy-encapsulin.

SEQUENCE LISTING

The nucleic and amino acid sequences listed in the accompanying sequence listing are shown using standard letter abbreviations for nucleotide bases, and three letter code for amino acids, as defined in 37 C.F.R. 1.822. Only one strand of each nucleic acid sequence is shown, but the complementary strand is understood as included by any reference to the displayed strand. The Sequence Listing is submitted as an ASCII text file in the form of the file named "Sequence.txt" (~1.5 MB), which was created on September 22, 2019, which is incorporated by reference herein.

DETAILED DESCRIPTION

As the HIV-1 pandemic continues to infect millions of people each year, the need for an effective vaccine increases. However, the development of such a vaccine has been stymied due to the difficulty in developing an immunogen capable of eliciting broadly neutralizing antibodies. The current disclosure meets these needs.

One of the major hurdles to the construction of an effective HIV-1 vaccine is focusing the immune response to regions of HIV proteins which mostly produce broadly neutralizing antibodies. As disclosed herein, a series of immunogens that elicit immune responses to the HIV-1 Env fusion peptide has been constructed. Such molecules have utility as both potential vaccines for HIV and as diagnostic molecules (for example, to detect and quantify target antibodies in a polyclonal serum response).

I. Summary of Terms

Unless otherwise noted, technical terms are used according to conventional usage. Definitions of common terms in molecular biology may be found in Benjamin Lewin, *Genes X*, published by Jones & Bartlett Publishers, 2009; and Meyers *et al.* (eds.), *The Encyclopedia of Cell Biology and Molecular Medicine*, published by Wiley-VCH in 16 volumes, 2008; and other similar references.

As used herein, the singular forms "a," "an," and "the," refer to both the singular as well as plural, unless the context clearly indicates otherwise. For example, the term "an antigen" includes single or plural antigens and can be considered equivalent to the phrase "at least one antigen." As used herein, the term "comprises" means "includes." It is further to be understood that any and all base sizes or amino acid sizes, and all molecular weight or molecular mass values, given for nucleic acids or polypeptides are approximate, and are provided for descriptive purposes, unless otherwise indicated. Although many methods and materials similar or equivalent to those described herein can be used, particular suitable methods and materials are described herein. In case of conflict, the present specification, including explanations of terms, will control. In addition, the materials, methods, and examples are illustrative only and not intended to be limiting. To facilitate review of the various embodiments, the following explanations of terms are provided:

Adjuvant: A vehicle used to enhance antigenicity. In some embodiments, an adjuvant can include a suspension of minerals (alum, aluminum hydroxide, or phosphate) on which antigen is adsorbed; or water-in-oil emulsion, for example, in which antigen solution is emulsified in mineral oil (Freund incomplete adjuvant), sometimes with the inclusion of killed mycobacteria (Freund's complete adjuvant) to further enhance antigenicity (inhibits degradation of antigen and/or causes influx of macrophages). In some embodiments, the adjuvant used in a disclosed immunogenic composition is a combination of lecithin and carbomer homopolymer (such as the ADJUPLEX™ adjuvant available from Advanced BioAdjuvants, LLC, see also Wegmann, Clin Vaccine Immunol, 22(9): 1004-1012, 2015). Additional adjuvants for use in the disclosed immunogenic compositions include the QS21 purified plant extract, Matrix M, AS01, MF59, and ALFQ adjuvants. Immunostimulatory oligonucleotides (such as those including a CpG motif) can also be used as adjuvants. Adjuvants include biological molecules (a "biological adjuvant"), such as costimulatory molecules. Exemplary adjuvants include IL-2, RANTES, GM-CSF, TNF-α, IFN-γ, G-CSF, LFA-3, CD72, B7-1, B7-2, OX-40L, 4-1BBL and toll-like receptor (TLR) agonists, such as TLR-9 agonists. Additional description of adjuvants can be found, for example, in Singh (ed.) Vaccine Adjuvants and Delivery Systems. Wiley-Interscience, 2007). Adjuvants can be used in combination with the disclosed immunogens.

Administration: The introduction of a composition into a subject by a chosen route. Administration can be local or systemic. For example, if the chosen route is intravenous, the composition (such as a composition including a disclosed immunogen) is administered by introducing the composition into a vein of the subject. Exemplary routes of administration include, but are not limited to, oral, injection (such as subcutaneous, intramuscular, intradermal, intraperitoneal, and intravenous), sublingual, rectal, transdermal (for example, topical), intranasal, vaginal, and inhalation routes.

Amino acid substitution: The replacement of an amino acid in a polypeptide with one or more different amino acids.

Antigen: A compound, composition, or substance that can stimulate the production of antibodies or a T cell response in an animal, including compositions that are injected or absorbed into an animal. An antigen reacts with the products of specific humoral or cellular immunity, including those induced by heterologous antigens, such as the disclosed HIV antigens. Examples of antigens include, but are not limited

to, polypeptides, peptides, lipids, polysaccharides, combinations thereof (such as glycopeptides) and nucleic acids containing antigenic determinants, such as those recognized by an immune cell. A vaccine antigen is an antigen that, when administered to a subject, elicits a prophylactic or therapeutic immune response in the subject.

Carrier protein: An immunogenic protein to which an antigen can be linked. When linked to a carrier, the antigen may become more immunogenic. Carriers are chosen to increase the immunogenicity of the antigen and/or to elicit antibodies against the carrier which are diagnostically, analytically, and/or therapeutically beneficial. Useful carrier proteins include polymeric carriers, which can be natural (for example, proteins from bacteria or viruses), semi-synthetic or synthetic materials containing one or more functional groups to which a reactant moiety can be attached.

Conjugated: A first moiety joined to a second moiety by a covalent bond. For example, a peptide (such as an HIV-1 Env fusion peptide) joined to a carrier (such as a self-assembling protein nanoparticle carrier as described herein) by a chemical linker (such as a Sulfo-SIAB linker).

Conservative variant: "Conservative" amino acid substitutions are those substitutions or deletions that do not substantially affect or decrease a function of a protein, such as the ability of the protein to elicit an immune response when administered to a subject. The term conservative amino acid substitution also includes the use of a substituted amino acid in place of an unsubstituted parent amino acid. Furthermore, individual substitutions, deletions or additions which alter, add or delete a single amino acid or a small percentage of amino acids (for instance less than 5%, in some embodiments less than 1%) in an encoded sequence are conservative variations where the alterations result in the substitution of an amino acid with a chemically similar amino acid.

The following six groups are examples of amino acids that are considered to be conservative substitutions for one another:

1) Alanine (A), Serine (S), Threonine (T);

2) Aspartic acid (D), Glutamic acid (E);

3) Asparagine (N), Glutamine (Q);

4) Arginine (R), Lysine (K);

5) Isoleucine (I), Leucine (L), Methionine (M), Valine (V); and

6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W).

Non-conservative substitutions are those that reduce an activity or function of the recombinant Env protein, such as the ability to elicit an immune response when administered to a subject. For instance, if an amino acid residue is essential for a function of the protein, even an otherwise conservative substitution may disrupt that activity. Thus, a conservative substitution does not alter the basic function of a protein of interest.

Consists essentially of and **Consists Of:** A polypeptide comprising an amino acid sequence that consists essentially of a specified amino acid sequence does not include any additional amino acid residues. However, the residues in the polypeptide can be modified to include non-peptide components, such as labels

(for example, fluorescent, radioactive, or solid particle labels), sugars or lipids, and the N- or C-terminus of the polypeptide can be joined (for example, by peptide bond) to heterologous amino acids, such as a cysteine (or other) residue in the context of a linker for conjugation chemistry. A polypeptide that consists of a specified amino acid sequence does not include any additional amino acid residues, nor does it include additional biological components, such as nucleic acids lipids, sugars, nor does it include labels. However, the N- or C-terminus of the polypeptide can be joined (for example, by peptide bond) to heterologous amino acids, such as a peptide tag, or a cysteine (or other) residue in the context of a linker for conjugation chemistry.

A polypeptide that consists or consists essentially of a specified amino acid sequence can be glycosylated or have an amide modification. A polypeptide that consists of or consists essentially of a particular amino acid sequence can be linked via its N- or C-terminus to a heterologous polypeptide, such as in the case of a fusion protein containing a first polypeptide consisting or a first sequence that is linked (via peptide bond) to a heterologous polypeptide consisting of a second sequence. In another example, the N- or C-terminus of a polypeptide that consists of or consists essentially of a particular amino acid sequence can be linked to a peptide linker (via peptide bond) that is further linked to one or more additional heterologous polypeptides. In a further example, the N- or C-terminus of a polypeptide that consists of or consists essentially of a particular amino acid sequence can be linked to one or more amino acid residues that facilitate further modification or manipulation of the polypeptide.

Control: A reference standard. In some embodiments, the control is a negative control sample obtained from a healthy patient. In other embodiments, the control is a positive control sample obtained from a patient diagnosed with HIV-1 infection. In still other embodiments, the control is a historical control or standard reference value or range of values (such as a previously tested control sample, such as a group of HIV-1 patients with known prognosis or outcome, or group of samples that represent baseline or normal values).

A difference between a test sample and a control can be an increase or conversely a decrease. The difference can be a qualitative difference or a quantitative difference, for example, a statistically significant difference. In some examples, a difference is an increase or decrease, relative to a control, of at least about 5%, such as at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 100%, at least about 150%, at least about 200%, at least about 250%, at least about 300%, at least about 350%, at least about 400%, at least about 500%, or greater than 500%.

Covalent bond: An interatomic bond between two atoms, characterized by the sharing of one or more pairs of electrons by the atoms. The terms "covalently bound" or "covalently linked" refer to making two separate molecules into one contiguous molecule. The terms include reference to conjugating an antigen (such as an HIV-1 Env fusion peptide) either directly or indirectly to a carrier molecule, for example indirectly with an intervening linker molecule.

Effective amount: An amount of agent, such as an immunogen, that is sufficient to generate a desired response, such as an immune response in a subject. It is understood that to obtain a protective immune response against an antigen of interest can require multiple administrations of a disclosed immunogen, and/or administration of a disclosed immunogen as the "prime" in a prime boost protocol wherein the boost immunogen can be different from the prime immunogen. Accordingly, an effective amount of a disclosed immunogen can be the amount of the immunogen sufficient to elicit a priming immune response in a subject that can be subsequently boosted with the same or a different immunogen to generate a protective immune response.

In one example, a desired response is to induce an immune response that inhibits or prevents HIV-1 infection. The HIV-1 infected cells do not need to be completely eliminated or prevented for the composition to be effective. For example, administration of an effective amount of the immunogen can induce an immune response that decreases the number of HIV-1 infected cells (or prevents the infection of cells) by a desired amount, for example, by at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 98%, or even at least 100% (elimination or prevention of detectable HIV-1 infected cells), as compared to the number of HIV-1 infected cells in the absence of the immunization.

Expression: Transcription or translation of a nucleic acid sequence. For example, a gene is expressed when its DNA is transcribed into an RNA or RNA fragment, which in some examples is processed to become mRNA. A gene may also be expressed when its mRNA is translated into an amino acid sequence, such as a protein or a protein fragment. In a particular example, a heterologous gene is expressed when it is transcribed into an RNA. In another example, a heterologous gene is expressed when its RNA is translated into an amino acid sequence. The term "expression" is used herein to denote either transcription or translation. Regulation of expression can include controls on transcription, translation, RNA transport and processing, degradation of intermediary molecules such as mRNA, or through activation, inactivation, compartmentalization or degradation of specific protein molecules after they are produced.

Expression Control Sequences: Nucleic acid sequences that regulate the expression of a heterologous nucleic acid sequence to which it is operatively linked. Expression control sequences are operatively linked to a nucleic acid sequence when the expression control sequences control and regulate the transcription and, as appropriate, translation of the nucleic acid sequence. Thus expression control sequences can include appropriate promoters, enhancers, transcription terminators, a start codon (ATG) in front of a protein-encoding gene, splicing signal for introns, maintenance of the correct reading frame of that gene to permit proper translation of mRNA, and stop codons. The term "control sequences" is intended to include, at a minimum, components whose presence can influence expression, and can also include additional components whose presence is advantageous, for example, leader sequences and fusion partner sequences. Expression control sequences can include a promoter.

A promoter is a minimal sequence sufficient to direct transcription. Also included are those promoter elements which are sufficient to render promoter-dependent gene expression controllable for cell-type specific, tissue-specific, or inducible by external signals or agents; such elements may be located in the

5' or 3' regions of the gene. Both constitutive and inducible promoters are included (see for example, Bitter *et al.*, *Methods in Enzymology* 153:516-544, 1987). For example, when cloning in bacterial systems, inducible promoters such as pL of bacteriophage lambda, plac, ptrp, ptac (ptrp-lac hybrid promoter) and the like may be used. In one embodiment, when cloning in mammalian cell systems, promoters derived from the genome of mammalian cells (such as metallothionein promoter) or from mammalian viruses (such as the retrovirus long terminal repeat; the adenovirus late promoter; the vaccinia virus 7.5K promoter) can be used. Promoters produced by recombinant DNA or synthetic techniques may also be used to provide for transcription of the nucleic acid sequences.

Fusion protein: A single polypeptide chain including the sequence of two or more heterologous proteins, often linked by a peptide linker. Reference to a first protein "fused" to a second protein indicates that the first and second proteins are contained within a single contiguous polypeptide chain. The first and second protein may be directly linked (for example, the C-terminus of the first protein is linked to the N-terminus of the second protein by a peptide bond), or indirectly linked (for example, the C-terminus of the first protein is directly linked to the N-terminus of a peptide linker by a peptide bond, and the C-terminus of the peptide linker is directly linked to the N-terminus of the second protein by a peptide bond).

Heterologous: Originating from a different genetic source.

Host cells: Cells in which a vector can be propagated and its DNA expressed. The cell may be prokaryotic or eukaryotic. The term also includes any progeny of the subject host cell. It is understood that all progeny may not be identical to the parental cell since there may be mutations that occur during replication. However, such progeny are included when the term "host cell" is used.

Human Immunodeficiency Virus Type 1 (HIV-1): A retrovirus that causes immunosuppression in humans (HIV-1 disease), and leads to a disease complex known as the acquired immunodeficiency syndrome (AIDS). "HIV-1 disease" refers to a well-recognized constellation of signs and symptoms (including the development of opportunistic infections) in persons who are infected by an HIV-1 virus, as determined by antibody or western blot studies. Laboratory findings associated with this disease include a progressive decline in T cells. Related viruses that are used as animal models include simian immunodeficiency virus (SIV), and feline immunodeficiency virus (FIV). Treatment of HIV-1 with HAART has been effective in reducing the viral burden and ameliorating the effects of HIV-1 infection in infected individuals.

HIV-1 envelope protein (Env): The HIV-1 Env protein is initially synthesized as a precursor protein of 845-870 amino acids in size. Individual precursor polypeptides form a homotrimer and undergo glycosylation within the Golgi apparatus as well as processing to remove the signal peptide, and cleavage by a cellular protease between approximately positions 511/512 to generate separate gp120 and gp41 polypeptide chains, which remain associated as gp120-gp41 protomers within the homotrimer. The ectodomain (that is, the extracellular portion) of the HIV-1 Env trimer undergoes several structural rearrangements from a prefusion mature (cleaved) closed conformation that evades antibody recognition, through intermediate conformations that bind to receptors CD4 and co-receptor (either CCR5 or CXCR4), to

a postfusion conformation. The HIV-1 Env ectodomain comprises the gp120 protein (approximately HIV-1 Env positions 31-511) and the gp41 ectodomain (approximately HIV-1 Env positions 512-644). An HIV-1 Env ectodomain trimer comprises a protein complex of three HIV-1 Env ectodomains. As used herein "HIV-1 Env ectodomain trimer" includes both soluble trimers (that is, trimers without gp41 transmembrane domain or cytoplasmic tail) and membrane anchored trimers (for example, trimers including a full-length gp41).

Mature gp120 includes approximately HIV-1 Env residues 31-511, contains most of the external, surface-exposed, domains of the HIV-1 Env trimer, and it is gp120 which binds both to cellular CD4 receptors and to cellular chemokine receptors (such as CCR5). A mature gp120 polypeptide is an extracellular polypeptide that interacts with the gp41 ectodomain to form an HIV-1 Env protomer that trimerizes to form the HIV-1 Env ectodomain trimer. The mature gp120 wild-type polypeptide is heavily N-glycosylated, giving rise to an apparent molecular weight of 120 kD. Native gp120 includes five conserved regions (C1-C5) and five regions of high variability (V1-V5).

Mature gp41 includes approximately HIV-1 Env residues 512-860, and includes cytosolic-, transmembrane-, and ecto-domains. The gp41 ectodomain (including approximately HIV-1 Env residues 512-644) can interact with gp120 to form an HIV-1 Env protomer that trimerizes to form the HIV-1 Env trimer. The HIV-1 Env fusion peptide is located at the N-terminus of gp41. Prior use of the HIV-1 Env fusion peptide for immunization (e.g., as described in Dingens et al, Plos Pathog., 14(7), e1007159, 2018; and Xu et al., Nat. Med., 24(6):857-867, 2018, each of which is incorporated by reference herein) illustrated HIV-1 Env fusion peptide-based immunization protocols.

The prefusion mature closed conformation of the HIV-1 Env ectodomain trimer is a structural conformation adopted by HIV-1 Env ectodomain trimer after cellular processing to a mature prefusion state with distinct gp120 and gp41 polypeptide chains, and before specific binding to the CD4 receptor. The three-dimensional structure of an exemplary HIV-1 Env ectodomain trimer in the prefusion mature closed conformation is known (see, *e.g.*, Pancera et al., Nature, 514:455-461, 2014). In the prefusion mature closed conformation, the HIV-1 Env ectodomain trimer includes a V1V2 domain "cap" at its membrane distal apex, with the V1V2 domain of each Env protomer in the trimer coming together at the membrane distal apex. At the membrane proximal aspect, the prefusion mature closed conformation of the HIV-1 Env ectodomain trimer includes distinct α6 and α7 helices. CD4 binding causes changes in the conformation of the HIV-1 Env ectodomain trimer, including disruption of the V1V1 domain cap, which "opens" as each V1V2 domain moves outward from the longitudinal axis of the Env trimer, and formation of the HR1 helix, which includes both the α6 and α7 helices (which are no longer distinct). These conformational changes bring the N-terminus of the fusion peptide within close proximity of the target cell membrane, and expose "CD4-induced" epitopes (such as the 17b epitope) that are present in the CD4-bound open conformation, but not the mature closed conformation, of the HIV-1 Env ectodomain trimer.

Unless context indicates otherwise, the numbering used in the disclosed HIV-1 Env proteins and fragments thereof (such as a gp120 and gp41) is relative to the HXB2 numbering scheme as set forth in

Numbering Positions in HIV Relative to HXB2CG Bette Korber *et al.*, Human Retroviruses and AIDS 1998: A Compilation and Analysis of Nucleic Acid and Amino Acid Sequences. Korber *et al.*, Eds. Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, NM, which is incorporated by reference herein in its entirety. For reference, the amino acid sequence of HIV-1 Env of *HXB2* is set forth as SEQ ID NO: 154 (GENBANK® GI:1906382, incorporated by reference herein as present in the database on June 20, 2014).

HXB2 (Clade B, SEQ ID NO: 13):
MRVKEKYQHLWRWGWRWGTMLLGMLMICSATEKLWVTVYYGVPVWKEATTTLFCASDAKAYDTEVHNVWATHACVPTDPN
PQEVVLVNVTENFNMWKNDMVEQMHEDIISLWDQSLKPCVKLTPLCVSLKCTDLKNDTNTNSSSGRMIMEKGEIKNCSFN
ISTSIRGKVQKEYAFFYKLDIIPIDNDTTSYKLTSCNTSVITQACPKVSFEPIPIHYCAPAGFAILKCNNKTFNGTGPCT
NVSTVQCTHGIRPVVSTQLLLNGSLAEEEVVIRSVNFTDNAKTIIVQLNTSVEINCTRPNNNTRKRIRIQRGPGRAFVTI
GKIGNMRQAHCNISRAKWNNTLKQIASKLREQFGNNKTIIFKQSSGGDPEIVTHSFNCGGEFFYCNSTQLFNSTWFNSTW
STEGSNNTEGSDTITLPCRIKQIINMWQKVGKAMYAPPISGQIRCSSNITGLLLTRDGGNSNNESEIFRPGGGDMRDNWR
SELYKYKVVKIEPLGVAPTKAKRRVVQREKRAVGIGALFLGFLGAAGSTMGAASMTLTVQARQLLSGIVQQQNNLLRAIE
AQQHLLQLTVWGIKQLQARILAVERYLKDQQLLGIWGCSGKLICTTAVPWNASWSNKSLEQIWNHTTWMEWDREINNYTS
LIHSLIEESQNQQEKNEQELLELDKWASLWNWFNITNWLWYIKLFIMIVGGLVGLRIVFAVLSIVNRVRQGYSPLSFQTH
LPTPRGPDRPEGIEEEGGERDRDRSIRLVNGSLALIWDDLRSLCLFSYHRLRDLLLIVTRIVELLGRRGWEALKYWWNLL
QYWSQELKNSAVSLLNATAIAVAEGTDRVIEVVQGACRAIRHIPRRIRQGLERILL

HIV-1 neutralizing antibody: An antibody that reduces the infectious titer of HIV-1 by binding to HIV-1 Env protein and inhibiting HIV-1 function. In some embodiments, neutralizing antibodies to HIV-1 can inhibit the infectivity of multiple strains of HIV-1, Teir-2 strain from multiple clades of HIV-1. In some embodiments, a disclosed immunogen can be administered to a subject to elicit an immune response that includes production of antibodies that specifically bind to the HIV-1 Env fusion peptide and neutralize Teir-2 strains of HIV-1 from multiple HIV-1 clades.

Immune response: A response of a cell of the immune system, such as a B cell, T cell, or monocyte, to a stimulus. In one embodiment, the response is specific for a particular antigen (an "antigen-specific response"). In one embodiment, an immune response is a T cell response, such as a CD4+ response or a CD8+ response. In another embodiment, the response is a B cell response, and results in the production of specific antibodies. "Priming an immune response" refers to treatment of a subject with a "prime" immunogen to induce an immune response that is subsequently "boosted" with a boost immunogen. Together, the prime and boost immunizations produce the desired immune response in the subject. "Enhancing an immune response" refers to co-administration of an adjuvant and an immunogenic agent, wherein the adjuvant increases the desired immune response to the immunogenic agent compared to administration of the immunogenic agent to the subject in the absence of the adjuvant.

Immunogen: A protein or a portion thereof that is capable of inducing an immune response in a mammal, such as a mammal infected or at risk of infection with a pathogen.

Immunogenic composition: A composition comprising a disclosed immunogen that elicits a measurable CTL response against the immunogen, or elicits a measurable B cell response (such as production of antibodies) against the immunogen, when administered to a subject. For *in vivo* use, the immunogenic composition will typically include the immunogen in a pharmaceutically acceptable carrier and may also include other agents, such as an adjuvant.

Immunogenic conjugate: A composition including of at least two heterologous molecules (such as an HIV-1 Env fusion peptide and a carrier, such as a self-assembling protein nanoparticle carrier) conjugated together. In a non-limiting example, a peptide (such as AVGIGAVF peptide, residues 1-8 of SEQ ID NO: 1) is linked to a protein carrier by a linker including a heterologous cysteine residue fused to the C-terminal residue of the peptide by peptide bond and a heterobifunctional moiety, wherein the heterobifunctional moiety is linked to a lysine residue on the carrier and the cysteine residue. In this example, the peptide is indirectly covalently linked to the carrier by the linker. Immunogenic conjugates are conjugates that are useful for eliciting a specific immune response to a molecule in the conjugate in a vertebrate. In some embodiments where the conjugate includes a viral antigen, the immune response is protective in that it enables the vertebrate animal to better resist infection from the virus from which the antigen is derived.

Inhibiting or treating a disease: Inhibiting the full development of a disease or condition, for example, in a subject who is at risk for a disease such as acquired immunodeficiency syndrome (AIDS). "Treatment" refers to a therapeutic intervention that ameliorates a sign or symptom of a disease or pathological condition after it has begun to develop. The term "ameliorating," with reference to a disease or pathological condition, refers to any observable beneficial effect of the treatment. Inhibiting a disease can include preventing or reducing the risk of the disease, such as preventing or reducing the risk of viral infection. The beneficial effect can be evidenced, for example, by a delayed onset of clinical symptoms of the disease in a susceptible subject, a reduction in severity of some or all clinical symptoms of the disease, a slower progression of the disease, a reduction in the viral load, an improvement in the overall health or well-being of the subject, or by other parameters that are specific to the particular disease. A "prophylactic" treatment is a treatment administered to a subject who does not exhibit signs of a disease or exhibits only early signs for the purpose of decreasing the risk of developing pathology.

Isolated: An "isolated" biological component has been substantially separated or purified away from other biological components, such as other biological components in which the component naturally occurs, such as other chromosomal and extrachromosomal DNA, RNA, and proteins. Proteins, peptides, nucleic acids, and viruses that have been "isolated" include those purified by standard purification methods. Isolated does not require absolute purity, and can include protein, peptide, nucleic acid, or virus molecules that are at least 50% isolated, such as at least 75%, 80%, 90%, 95%, 98%, 99%, or even 99.9% isolated.

Linked: The term "linked" means joined together, either directly or indirectly. For example, a first moiety may be covalently or noncovalently (e.g., electrostatically) linked to a second moiety. This includes, but is not limited to, covalently bonding one molecule to another molecule, noncovalently bonding one molecule to another (e.g. electrostatically bonding), non-covalently bonding one molecule to another molecule by hydrogen bonding, non-covalently bonding one molecule to another molecule by van der Waals forces, and any and all combinations of such couplings. Indirect attachment is possible, such as by using a "linker". In several embodiments, linked components are associated in a chemical or physical manner so that the components are not freely dispersible from one another, at least until contacting a cell, such as an immune cell.

Linker: One or more molecules or groups of atoms positioned between two moieties. Typically, linkers are bifunctional, *i.e.*, the linker includes a functional group at each end, wherein the functional groups are used to couple the linker to the two moieties. The two functional groups may be the same, *i.e.*, a homobifunctional linker, or different, *i.e.*, a heterobifunctional linker. In several embodiments, a peptide linker can be used to link the C-terminus of a first protein to the N-terminus of a second protein. Non-limiting examples of peptide linkers include glycine-serine peptide linkers, which are typically not more than 10 amino acids in length. In a non-limiting example, a peptide (such as AVGIGAVF peptide, residues 1-8 of SEQ ID NO: 1) is linked to a protein carrier by a linker including a heterologous cysteine residue fused to the C-terminal residue of the peptide by peptide bond and a heterobifunctional moiety, wherein the heterobifunctional moiety is linked to a lysine residue on the carrier and the cysteine residue.

Nucleic acid molecule: A polymeric form of nucleotides, which may include both sense and anti-sense strands of RNA, cDNA, genomic DNA, and synthetic forms and mixed polymers of the above. A nucleotide refers to a ribonucleotide, deoxynucleotide or a modified form of either type of nucleotide. The term "nucleic acid molecule" as used herein is synonymous with "nucleic acid" and "polynucleotide." A nucleic acid molecule is usually at least 10 bases in length, unless otherwise specified. The term includes single- and double-stranded forms of DNA. A polynucleotide may include either or both naturally occurring and modified nucleotides linked together by naturally occurring and/or non-naturally occurring nucleotide linkages. "cDNA" refers to a DNA that is complementary or identical to an mRNA, in either single stranded or double stranded form. "Encoding" refers to the inherent property of specific sequences of nucleotides in a polynucleotide, such as a gene, a cDNA, or an mRNA, to serve as templates for synthesis of other polymers and macromolecules in biological processes having either a defined sequence of nucleotides (i.e., rRNA, tRNA and mRNA) or a defined sequence of amino acids and the biological properties resulting therefrom.

Pattern recognition receptor: A protein receptor expressed by cells of the immune system to identify pathogen-associated molecular patterns (PAMPS) as well as damage associated molecular patterns (DAMPs). PAMP or DAMP activation of pattern recognition receptors induces an intracellular signaling cascade resulting in the alteration of the host cell's transcription profile to induce expression of pro-inflammatory and pro-survival genes that enhance adaptive immunity. Non-limiting examples of pattern recognition receptors (PRRs) include Toll-like receptors (TLR), Stimulator of Interferon Genes receptor (STING), C-type lectin receptors (CLR), RIG-I-like receptors (RLR), and NOD-like receptors (NLR). In some embodiments, agonists of such pattern recognition receptors can be linked to a disclosed immunogenic conjugate to target the conjugate to pattern recognition receptor expressing cells (i.e., cells of the immune system) to enhance the immune response to the immunogenic conjugate.

Toll-like receptors (TLRs) 1-13 are transmembrane PRRs that recognize a diverse range of PAMPs. TLRs can be divided into two broad categories - those that are localized to the cell surface and those that are localized to the endosomal lumen. TLRs that are present on the cell surface are important in recognition of bacterial pathogens. TLRs that are localized to the lumen of endosomes, such as TLRs 3, 7, 8, and 9, serve to recognize nucleic acids and are thus thought to be important in the promotion of antiviral immune

responses. TLR-7 and TLR-8 recognize ssRNA. Several different imidazoquinoline compounds are known TLR-7/8 agonists. TLR-9 recognizes unmethylated deoxycytidylate-phosphate-deoxyguanylate (CpG) DNA, found primarily in bacteria.

The NOD-like receptors (NLRs) and the RIG-I-like receptors (RLRs) are localized to the cytoplasm. Non-limiting examples of RLRs include RIG-I, MDA5, and LGP2. There are 22 human NLRs that can be subdivided into the five structurally related NLR families A, B, C, P, and X. All NLRs have three domains: an N-terminal domain involved in signaling, a nucleotide-binding NOD domain, and a C-terminal leucine rich region (LRR) important for ligand recognition. Non-limiting examples of NLRs include NALP3 and NOD2.

For more information on pattern recognition receptors, see Wales et al., Biochem Soc Trans., 35:1501-1503, 2007.

Pharmaceutically acceptable carriers: The pharmaceutically acceptable carriers of use are conventional. *Remington's Pharmaceutical Sciences*, by E. W. Martin, Mack Publishing Co., Easton, PA, 19th Edition, 1995, describes compositions and formulations suitable for pharmaceutical delivery of the disclosed immunogens.

In general, the nature of the carrier will depend on the particular mode of administration being employed. For instance, parenteral formulations usually comprise injectable fluids that include pharmaceutically and physiologically acceptable fluids such as water, physiological saline, balanced salt solutions, aqueous dextrose, glycerol or the like as a vehicle. For solid compositions (*e.g.*, powder, pill, tablet, or capsule forms), conventional non-toxic solid carriers can include, for example, pharmaceutical grades of mannitol, lactose, starch, or magnesium stearate. In addition to biologically neutral carriers, pharmaceutical compositions to be administered can contain minor amounts of non-toxic auxiliary substances, such as wetting or emulsifying agents, preservatives, and pH buffering agents and the like, for example, sodium acetate or sorbitan monolaurate. In particular embodiments, suitable for administration to a subject the carrier may be sterile, and/or suspended or otherwise contained in a unit dosage form containing one or more measured doses of the composition suitable to elicit the desired anti-HIV-1 immune response. It may also be accompanied by medications for its use for treatment purposes. The unit dosage form may be, for example, in a sealed vial that contains sterile contents or a syringe for injection into a subject, or lyophilized for subsequent solubilization and administration or in a solid or controlled release dosage.

Polypeptide: Any chain of amino acids, regardless of length or post-translational modification (*e.g.*, glycosylation or phosphorylation). "Polypeptide" applies to amino acid polymers including naturally occurring amino acid polymers and non-naturally occurring amino acid polymer as well as in which one or more amino acid residue is a non-natural amino acid, for example, an artificial chemical mimetic of a corresponding naturally occurring amino acid. A "residue" refers to an amino acid or amino acid mimetic incorporated in a polypeptide by an amide bond or amide bond mimetic. A polypeptide has an amino

terminal (N-terminal) end and a carboxy terminal (C-terminal) end. "Polypeptide" is used interchangeably with peptide or protein, and is used herein to refer to a polymer of amino acid residues.

Prime-boost immunization: An immunotherapy including administration of multiple immunogens over a period of time to elicit the desired immune response.

Recombinant: A recombinant nucleic acid molecule is one that has a sequence that is not naturally occurring, for example, includes one or more nucleic acid substitutions, deletions or insertions, and/or has a sequence that is made by an artificial combination of two otherwise separated segments of sequence. This artificial combination can be accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, for example, by genetic engineering techniques.

A recombinant virus is one that includes a genome that includes a recombinant nucleic acid molecule.

A recombinant protein is one that has a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two otherwise separated segments of sequence. In several embodiments, a recombinant protein is encoded by a heterologous (for example, recombinant) nucleic acid that has been introduced into a host cell, such as a bacterial or eukaryotic cell, or into the genome of a recombinant virus.

Self-assembling protein nanoparticle: A multi-subunit protein-based nanoparticle formed from subunit monomers that self-assemble under suitable conditions to form the nanoparticle (typically globular in shape). Non-limiting examples of self-assembling protein nanoparticles include ferritin nanoparticles (see, *e.g.*, Zhang, *Y. Int. J. Mol. Sci.*, 12:5406-5421, 2011, incorporated by reference herein), encapsulin nanoparticles (see, *e.g.*, Sutter *et al.*, Nature Struct. and Mol. Biol., 15:939-947, 2008, incorporated by reference herein), Sulfur Oxygenase Reductase (SOR) nanoparticles (see, *e.g.*, Urich *et al.*, Science, 311:996-1000, 2006, incorporated by reference herein), lumazine synthase nanoparticles (see, *e.g.*, Zhang *et al.*, *J. Mol. Biol.*, 306: 1099-1114, 2001), and pyruvate dehydrogenase nanoparticles (see, *e.g.*, Izard et al., PNAS 96: 1240-1245, 1999, incorporated by reference herein). Ferritin, encapsulin, SOR, lumazine synthase, and pyruvate dehydrogenase are monomeric proteins that self-assemble into a globular protein complexes that in some cases consists of 24, 60, 24, 60, and 60 protein subunits, respectively. In some examples, ferritin, encapsulin, SOR, lumazine synthase, or pyruvate dehydrogenase subunits are fused to a disclosed heterologous carrier protein (such as an rTT, CRM197, or HiD carrier protein) and self-assembled into a protein nanoparticle presenting the carrier protein, which can subsequently be conjugated to HIV-1 Env fusion proteins to generate an immunogenic conjugate to elicit or prime an immune response to HIV-1 Env in a subject.

Sequence identity: The similarity between amino acid sequences is expressed in terms of the similarity between the sequences, otherwise referred to as sequence identity. Sequence identity is frequently measured in terms of percentage identity; the higher the percentage, the more similar the two sequences are. Homologs, orthologs, or variants of a polypeptide will possess a relatively high degree of sequence identity when aligned using standard methods.

Methods of alignment of sequences for comparison are well known in the art. Various programs and alignment algorithms are described in: Smith & Waterman, *Adv. Appl. Math.* 2:482, 1981; Needleman & Wunsch, *J. Mol. Biol.* 48:443, 1970; Pearson & Lipman, *Proc. Natl. Acad. Sci. USA* 85:2444, 1988; Higgins & Sharp, *Gene*, 73:237-44, 1988; Higgins & Sharp, *CABIOS* 5:151-3, 1989; Corpet *et al.*, *Nuc. Acids Res.* 16:10881-90, 1988; Huang *et al. Computer Appls. In the Biosciences* 8, 155-65, 1992; and Pearson *et al.*, *Meth. Mol. Bio.* 24:307-31, 1994. Altschul *et al.*, *J. Mol. Biol.* 215:403-10, 1990, presents a detailed consideration of sequence alignment methods and homology calculations.

Variants of a polypeptide are typically characterized by possession of at least about 75%, for example, at least about 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity counted over the full length alignment with the amino acid sequence of interest. Proteins with even greater similarity to the reference sequences will show increasing percentage identities when assessed by this method, such as at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% sequence identity. When less than the entire sequence is being compared for sequence identity, homologs and variants will typically possess at least 80% sequence identity over short windows of 10-20 amino acids, and may possess sequence identities of at least 85% or at least 90% or 95% depending on their similarity to the reference sequence. Methods for determining sequence identity over such short windows are available at the NCBI website on the internet.

As used herein, reference to "at least 90% identity" (or similar language) refers to "at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or even 100% identity" to a specified reference sequence.

Signal Peptide: A short amino acid sequence (*e.g.*, approximately 18-30 amino acids in length) that directs newly synthesized secretory or membrane proteins to and through membranes (for example, the endoplasmic reticulum membrane). Signal peptides are typically located at the N-terminus of a polypeptide and are removed by signal peptidases after the polypeptide has crossed the membrane. Signal peptide sequences typically contain three common structural features: an N-terminal polar basic region (n-region), a hydrophobic core, and a hydrophilic c-region). An exemplary signal peptide sequence is set forth as MDSKGSSQKGSRLLLLLVVSNLLLPQGVVA (SEQ ID NO: 220).

Specifically bind: When referring to the formation of an antibody:antigen protein complex, or a protein:protein complex, refers to a binding reaction which determines the presence of a target protein, peptide, or polysaccharide (for example, a glycoprotein), in the presence of a heterogeneous population of proteins and other biologics. Thus, under designated conditions, a particular antibody or protein binds preferentially to a particular target protein, peptide or polysaccharide (such as an antigen present on the surface of a pathogen, for example, gp120) and does not bind in a significant amount to other proteins or polysaccharides present in the sample or subject. Specific binding can be determined by standard methods. A first protein or antibody specifically binds to a target protein when the interaction has a K_D of less than 10^{-6} Molar, such as less than 10^{-7} Molar, less than 10^{-8} Molar, less than 10^{-9}, or even less than 10^{-10} Molar.

Subject: Living multicellular vertebrate organisms, a category that includes human and non-human mammals. In an example, a subject is a human. In a particular example, the subject is a newborn infant. In an additional example, a subject is selected that is in need of inhibiting of an HIV-1 infection. For example, the subject is either uninfected and at risk of HIV-1 infection or is infected in need of treatment.

Under conditions sufficient for: A phrase that is used to describe any environment that permits a desired activity.

Vaccine: A pharmaceutical composition that elicits a prophylactic or therapeutic immune response in a subject. In some cases, the immune response is a protective immune response. Typically, a vaccine elicits an antigen-specific immune response to an antigen of a pathogen, for example a viral pathogen, or to a cellular constituent correlated with a pathological condition. A vaccine may include a polynucleotide (such as a nucleic acid encoding a disclosed antigen), a peptide or polypeptide (such as a disclosed antigen), a virus, a cell or one or more cellular constituents. In one specific, non-limiting example, a vaccine reduces the severity of the symptoms associated with HIV-1 infection and/or decreases the viral load compared to a control. In another non-limiting example, a vaccine reduces HIV-1 infection compared to a control.

VRC34: An antibody that binds to the fusion peptide of HIV-1 any neutralizing HIV-1 infection. VRC34. Unless context indicates otherwise, "VRC34" refers to the VRC34.01 antibody disclosed by Kong et al. (Science, 352, 828-833, 2016). Sequences of the heavy and light chain variable regions of the VRC34.01 antibody are available, for example, as GenBank Accession Nos. ANF29805.1 and ANF29798.1, respectively, each of which is incorporated by reference herein. The VRC34 antibody can be used to assess the antigenicity fo the disclosed immunogenic conjugates of HIV-1 Env fusion peptides conjugated to a self-assembling protein nanoparticle carrier.

II. Immunogenic Conjugates

Immunogenic conjugates are provided herein that include HIV-1 Env fusion peptides conjugated to a self-assembling protein nanoparticle carrier. In several embodiments, the immunogenic conjugates can be used to generate a neutralizing immune response to HIV-1 in a subject, for example, to treat or prevent an HIV-1 infection in the subject. The immunogenic conjugate provides a multivalent platform with superior binding capability for engaging HIV-1 Env fusion peptide-directed broadly neutralizing antibodies and can be used, for example, to prime an immune response in a subject that targets the HIV-1 Env fusion peptide epitope. The components of the immunogenic conjugate are discussed in more detail below.

A. Self-Assembling Protein Nanoparticle Carrier

The immunogenic conjugates provided herein include HIV-1 Env fusion peptides conjugated to a self-assembling protein nanoparticle carrier. The self-assembling protein nanoparticle carrier is formed from a multimer of fusion proteins that each include a self-assembling protein nanoparticle subunit fused to a heterologous carrier protein. The subunit and the carrier protein can be directly fused (the C-terminus of one is linked by peptide bond to the N-terminus of the other) or indirectly fused via a peptide linker. Following

expression of the fusion proteins (typically in a cellular system where the fusion proteins are secreted into the supernatant), the self-assembling protein nanoparticle subunits of the fusion proteins self-assemble under suitable conditions into the protein nanoparticle, forming a protein complex containing the protein nanoparticle and displaying the heterologous carrier proteins (at least one of which is fused to each self-assembling protein nanoparticle subunit). This protein complex is the self-assembling protein nanoparticle carrier to which HIV-1 Env fusion peptides are conjugated to form an immunogenic conjugate of the present disclosure.

1. Self-assembling protein nanoparticle subunit

The fusion proteins of the self-assembling protein nanoparticle carrier include a self-assembling protein nanoparticle subunit fused to a heterologous carrier protein. The self-assembling protein nanoparticle subunit is a monomer of a self-assembling protein nanoparticle, or a fragment of such a monomer that retains the portion of the monomer required for self-assembly. Non-limiting examples of self-assembling protein nanoparticle subunits that can be included in the fusion protein to form a self-assembling protein nanoparticle carrier include lumazine synthase nanoparticle subunits, ferritin nanoparticle subunts, encapsulin nanoparticle subunits, Sulfur Oxygenase Reductase (SOR) nanoparticle subunits, Bacteriophage Q Beta Capsid protein (qbeta) subunits, Dihydrolipoyl transacetylase protein (e2p) subunits, Phosphopantetheine Adenylyltransferase (6ccq) subunits, Glutamate Synthase (1f52) subunits, Calcium/calmodulin dependent protein kinase IIa (CaMKIIa), C-terminal fragment (5U6Y) subunits, HIV capsid oligomerization domain subunits, Hexamer subunits, and T4 fibritin Foldon domain (Fd) subunits. In a preferred embodiment, the self-assembling protein nanoparticle subunit included in the fusion protein is a ferritin subunit. In another preferred embodiment, the self-assembling protein nanoparticle subunit included in the fusion protein is a lumazine synthase subunit.

a. Ferritin

In some embodiments, any of the disclosed heterologous carrier proteins (such as an rTT, CRM197, or HiD carrier protein) can be linked to a ferritin subunit to construct a self-assembling ferritin nanoparticle carrier including a ferritin nanoparticle fused to a plurality of the heterologous carrier proteins. Ferritin nanoparticles and their use for immunization purposes (e.g., for immunization against influenza antigens) have been disclosed, for example, in Kanekiyo et al. (Nature, 499:102-106, 2013, incorporated by reference herein in its entirety). Ferritin is a globular protein that is found in all animals, bacteria, and plants, and which acts primarily to control the rate and location of polynuclear $Fe(III)_2O_3$ formation through the transportation of hydrated iron ions and protons to and from a mineralized core. Ferritin nanoparticles are formed from 24 copies of the ferritin subunit. The globular form of the ferritin nanoparticle is made up of monomeric subunits. Non-limiting examples of the sequence of self-assembling ferritin subunits for use in the embodiments provided herein include:

DIEKLLNEQVNKEMQSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKFEGLT
QIFQKAYEHEQHISESINNIVDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDKIELIGNENHGLYLADQYVKGIAKS
RKS (SEQ ID NO: 14)

DIIKLLNEQVNKEMQSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPHHKFHGLT
HIFHKAYHHEQHISESINNIVDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDKIELIGNENHGLYLADQYVKGIAKS
RKS (SEQ ID NO: 15)

DIIKLLNEQVNKEMNSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKFEGLT
QIFQKAYEHEQHISESINNIVDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDKIELIGNENHGLYLADQYVKGIAKS
RKS (SEQ ID NO: 16)

DIIKLLNEQVNKEMQSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKFEGLT
QIFQKAYEHEQHISESINNIVDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDKIELIGNENHGLYLADQYVKGIAKS
RKS (SEQ ID NO: 17)

DIIKLLNEQVNKEMDSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKFEGLT
QIFQKAYEHEQHISESINNIVDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDKIELIGNENHGLYLADQYVKGIAKS
RKS (SEQ ID NO: 18)

and the following two sequences which include C-terminal truncations:

DIIKLLNEQVNKEMNSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKFEGLT
QIFQKAYEHEQHISESINNIVDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDKIELIG (SEQ ID NO: 19)

DIIKLLNEQVNKEMNSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKFEGLT
QIFQKAYEHEQHISESINNIVDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDKIELIGNEN (SEQ ID NO: 20)

Additional ferritin subunits are provided with one or more cysteine substitutions to introduce non-native disulfide bond(s) that stabilize the ferritin nanoparticle formed from the self-assembled subunits. As used herein, a non-native disulfide bond introduced into a self-assembling protein nanoparticle subunit that "stabilizes" the nanoparticle formed from oligomerization of the subunit increases retention of the assembled nanoparticle compared to a control nanoparticle formed from subunits lacking the disulfide bond. The "stabilization" of the nanoparticle can be, for example, an increase in resistance to disassembly of the subunits compared to a corresponding native subunit sequence. Non-limiting examples of ferritin subunits are provided with one or more cysteine substitutions to introduce non-native disulfide bond(s) that stabilize the ferritin nanoparticle formed from the self-assembled subunits include:

Ferr_Hp_DS01
MLSKDIIKLLNEQVNKEMQSSNLYMSMSCWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIFLNENNVPCQLTSISAPEHKF
EGLTQIFQKAYEHEQHISESINNIVDHAIKSKDHCTFNFLQWYVAEQCEEEVLFKDILDKIELIGNENHGLYLADQYVKG
IAKSRKS (SEQ ID NO: 258)

Ferr_Hp_DS02
MLSKDIIKLLNEQVNKEMQSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQLTGCISAPEHK
FEGLTQIFQKAYEHEQHISESINNIVDHAIKSKDHCTFNFLQWYVAEQCEEEVLFKDILDKIELIGNENHGLYLADQYVK
GIAKSRKS (SEQ ID NO: 259)

Ferr_Hp_DS03
MLSKDIIKLLNEQVNKEMQSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQLTCISCPEHKF
EGLTQIFQKAYEHEQHISESINNIVDHAIKSKDHCTFNFLQWYVAEQCEEEVLFKDILDKIELIGNENHGLYLADQYVKG
IAKSRKS (SEQ ID NO: 260)

Ferr_Hp_DS04

MLSKDIIKLLNEQVNKEMQSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQLTSCSAPEHKF
EGLTQIFQKAYEHEQHISESINNIVDHAIKSKDHCTFNFLQWYVAEQCEEEVLFKDILDKIELIGNENHGLYLADQYVKG
IAKSRKS (SEQ ID NO: 261)

Ferr_Hp_DS05
MLSKDIIKLLNEQVNKEMQSSNLYMSMSCWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIFLNCNNVPCQLTSISAPEHKF
EGLTQIFQKAYEHEQHISESINNIVDHAIKSKDHATFNFLQWYVAEQHEEECLFKDILDKIELIGNENHGLYLADQYVKG
IAKSRKS (SEQ ID NO: 262)

Ferr_Hp_DS06
MLSKDIIKLLNEQVNKEMQSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIFLNCNNVPVQLTGCISAPEHK
FEGLTQIFQKAYEHEQHISESINNIVDHAIKSKDHATFNFLQWYVAEQHEEECLFKDILDKIELIGNENHGLYLADQYVK
GIAKSRKS (SEQ ID NO: 263)

Ferr_Hp_DS07
MLSKDIIKLLNEQVNKEMQSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIFLNCNNVPVQLTCISCPEHKF
EGLTQIFQKAYEHEQHISESINNIVDHAIKSKDHATFNFLQWYVAEQHEEECLFKDILDKIELIGNENHGLYLADQYVKG
IAKSRKS (SEQ ID NO: 264)

Ferr_Hp_DS08
MLSKDIIKLLNEQVNKEMQSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIFLNCNNVPVQLTSCSAPEHKF
EGLTQIFQKAYEHEQHISESINNIVDHAIKSKDHATFNFLQWYVAEQHEEECLFKDILDKIELIGNENHGLYLADQYVKG
IAKSRKS (SEQ ID NO: 265)

Ferr_Hp_DS09
MLSKDIIKLLNEQVNKEMQSSNLYMSMSCWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIFLNENNVPCQLTSISAPEHKF
EGLTQIFQKAYEHEQHISESINNICDHAIKSKDCATFNFLQWYVAEQHEEEVLFKDILDKIELIGNENHGLYLADQYVKG
IAKSRKS (SEQ ID NO: 266)

Ferr_Hp_DS10
MLSKDIIKLLNEQVNKEMQSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQLTGCISAPEHK
FEGLTQIFQKAYEHEQHISESINNICDHAIKSKDCATFNFLQWYVAEQHEEEVLFKDILDKIELIGNENHGLYLADQYVK
GIAKSRKS (SEQ ID NO: 267)

Ferr_Hp_DS11
MLSKDIIKLLNEQVNKEMQSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQLTCISCPEHKF
EGLTQIFQKAYEHEQHISESINNICDHAIKSKDCATFNFLQWYVAEQHEEEVLFKDILDKIELIGNENHGLYLADQYVKG
IAKSRKS (SEQ ID NO: 268)

Ferr_Hp_DS12
MLSKDIIKLLNEQVNKEMQSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQLTSCSAPEHKF
EGLTQIFQKAYEHEQHISESINNICDHAIKSKDCATFNFLQWYVAEQHEEEVLFKDILDKIELIGNENHGLYLADQYVKG
IAKSRKS (SEQ ID NO: 269)

Ferr_pf_DS01
MLSERMLKALNDQLNRELYSAYLYFAMACYFEDLGLEGFANWMKAQAEEEIGHALRFYNYIYCRNGRCELDEIPKPPKEW
ESPLKAFEAAYEHEKFISKSIYELAALAEEEKDYSTRAFLEWFINEQVEEECSVKKILDKLKFAKDSPQILFMLDKELSA
RAPKLPGLLMQGGE (SEQ ID NO: 270)

Ferr_pf_DS02
MLSERMLKALNDQLNRELYSAYLYFAMAAYFEDLGLEGFANWMKAQAEEEIGHALRFYNYIYCRNGRVELDCIPCPPKEW
ESPLKAFEAAYEHEKFISKSIYELAALAEEEKDYSTRAFLEWFINEQVEEECSVKKILDKLKFAKDSPQILFMLDKELSA
RAPKLPGLLMQGGE (SEQ ID NO: 271)

Ferr_pf_DS03
MLSERMLKALNDQLNRELYSAYLYFAMAAYFEDLGLEGFANWMKAQAEEEIGHALRFYNYIYCRNGRVELDECPKPPKEW
ESPLKAFEAAYEHEKFISKSIYELAALAEEEKDYSTRAFLEWFINEQVEEECSVKKILDKLKFAKDSPQILFMLDKELSA
RAPKLPGLLMQGGE (SEQ ID NO: 272)

Ferr_pf_DS04

MLSERMLKALNDQLNRELYSAYLYFAMAAYFEDLGLEGFANWMKAQAEEEIGHALRFYNYIYCRNGRVELDCIPKPPKEW
ESPLKAFEAAYEHEKFISKSIYELAALAEEEKDYSTRAFLEWFINEQVEEECSVKKILDKLKFAKDSPQILFMLDKELSA
RAPKLPGLLMQGGE (SEQ ID NO: 273)

Ferr_pf_DS05
MLSERMLKALNDQLNRELYSAYLYFAMAAYFEDLGLEGFANWMKAQAEEEIGHALRFYNYIYCRNGRVELDGCIPKPPKE
WESPLKAFEAAYEHEKFISKSIYELAALAEEEKDYSTRAFLEWFINEQVEEECSVKKILDKLKFAKDSPQILFMLDKELS
ARAPKLPGLLMQGGE (SEQ ID NO: 274)

Ferr_pf_DS06
MLSERMLKALNDQLNRELYSAYLYFAMAAYFEDLGLEGFANWMKAQAEEEIGHALRFYNYIYCRNGRVELDEIPKPPKEW
ESPLKAFEAAYEHEKFISKSIYELAALAEEEKDYSTRAFLEWFINEQVEEECSVKKILDKLKFAKDSPQILFMLDKELSA
RAPKLPGGGC (SEQ ID NO: 275)

Ferr_pf_DS07
MLSERMLKALNDQLNRELYSAYLYFAMAAYFEDLGLEGFANWMKAQAEEEIGHALRFYNYIYCRNGRVELDEIPKPPKEW
ESPLKAFEAAYEHEKFISKSIYELAALAEEEKDYSTRAFLEWFINEQVEEECSVKKILDKLKFAKDSPQILFMLDKELSA
RAPKLPGGGWC (SEQ ID NO: 276)

Ferr_pf_DS08
MLSERMLKALNDQLNRELYSAYLYFAMACYFEDLGLEGFANWMKAQAEEEIGHALRFYNYIYDRNGRCELDEIPKPPKEW
ESPLKAFEAAYEHEKFISKSIYELAALAEEEKDCSTRAFLEWFCNEQVEEEASVKKILDKLKFAKDSPQILFMLDKELSA
RAPKLPGLLMQGGE (SEQ ID NO: 277)

Ferr_pf_DS09
MLSERMLKALNDQLNRELYSAYLYFAMAAYFEDLGLEGFANWMKAQAEEEIGHALRFYNYIYDRNGRVELDCIPCPPKEW
ESPLKAFEAAYEHEKFISKSIYELAALAEEEKDCSTRAFLEWFCNEQVEEEASVKKILDKLKFAKDSPQILFMLDKELSA
RAPKLPGLLMQGGE (SEQ ID NO: 278)

Ferr_pf_DS10
MLSERMLKALNDQLNRELYSAYLYFAMAAYFEDLGLEGFANWMKAQAEEEIGHALRFYNYIYDRNGRVELDECPKPPKEW
ESPLKAFEAAYEHEKFISKSIYELAALAEEEKDCSTRAFLEWFCNEQVEEEASVKKILDKLKFAKDSPQILFMLDKELSA
RAPKLPGLLMQGGE (SEQ ID NO: 279)

Ferr_pf_DS11
MLSERMLKALNDQLNRELYSAYLYFAMAAYFEDLGLEGFANWMKAQAEEEIGHALRFYNYIYDRNGRVELDCIPKPPKEW
ESPLKAFEAAYEHEKFISKSIYELAALAEEEKDCSTRAFLEWFCNEQVEEEASVKKILDKLKFAKDSPQILFMLDKELSA
RAPKLPGLLMQGGE (SEQ ID NO: 280)

Ferr_pf_DS12
MLSERMLKALNDQLNRELYSAYLYFAMAAYFEDLGLEGFANWMKAQAEEEIGHALRFYNYIYDRNGRVELDGCIPKPPKE
WESPLKAFEAAYEHEKFISKSIYELAALAEEEKDCSTRAFLEWFCNEQVEEEASVKKILDKLKFAKDSPQILFMLDKELS
ARAPKLPGLLMQGGE (SEQ ID NO: 281)

Ferr_pf_DS13
MLSERMLKALNDQLNRELYSAYLYFAMAAYFEDLGLEGFANWMKAQAEEEIGHALRFYNYIYDRNGRVELDEIPKPPKEW
ESPLKAFEAAYEHEKFISKSIYELAALAEEEKDCSTRAFLEWFCNEQVEEEASVKKILDKLKFAKDSPQILFMLDKELSA
RAPKLPGGGC (SEQ ID NO: 282)

Ferr_pf_DS14
MLSERMLKALNDQLNRELYSAYLYFAMAAYFEDLGLEGFANWMKAQAEEEIGHALRFYNYIYDRNGRVELDEIPKPPKEW
ESPLKAFEAAYEHEKFISKSIYELAALAEEEKDCSTRAFLEWFCNEQVEEEASVKKILDKLKFAKDSPQILFMLDKELSA
RAPKLPGGGWC (SEQ ID NO: 283)

Ferr_Mt_DS01
MTEYEGPKTKFHALMQEQIHNEFTAAQQYVCIAVYFDSEDLPQLAKHFYSQAVEERNHAMMLVQHLLCRDLRVECPGVDT
VRNQFDRPREALALALDQERTVTDQVGRLTAVARDEGDFLGEQFMQWFLQEQIEEVCLMATLVRVADRAGANLFELENFV
AREVDVAPAASGAPHAAGGRL (SEQ ID NO: 284)

Ferr_Mt_DS02

MTEYEGPKTKFHALMQEQIHNEFTAAQQYVAIAVYFDSEDLPQLAKHFYSQAVEERNHAMMLVQHLLCRDLRVEIPGCDT
VRNQFDRPREALALALDQERTVTDQVGRLTAVARDEGDFLGEQFMQWFLQEQIEEVCLMATLVRVADRAGANLFELENFV
AREVDVAPAASGAPHAAGGRL (SEQ ID NO: 285)

Ferr_Mt_DS03
MTEYEGPKTKFHALMQEQIHNEFTAAQQYVAIAVYFDSEDLPQLAKHFYSQAVEERNHAMMLVQHLLCRDLRVCIPGVDT
VRCQFDRPREALALALDQERTVTDQVGRLTAVARDEGDFLGEQFMQWFLQEQIEEVCLMATLVRVADRAGANLFELENFV
AREVDVAPAASGAPHAAGGRL (SEQ ID NO: 286)

Ferr_Mt_DS04
MTEYEGPKTKFHALMQEQIHNEFTAAQQYVAIAVYFDSEDLPQLAKHFYSQAVEERNHAMMLVQHLLCRDLRVEIPGVDT
VRNQFDRPREALALALDQERTVTDQVGRLTAVARDEGDFLGEQFMQWFLQEQIEEVCLMATLVRVADRAGANLFELENFV
AREVDVAPAASGAPHAAGGRC (SEQ ID NO: 287)

Ferr_Mt_DS05
MTEYEGPKTKFHALMQEQIHNEFTAAQQYVAIAVYFDSEDLPQLAKHFYSQAVEERNHAMMLVQHLLCRDLRVEIPGVDT
VRNQFDRPREALALALDQERTVTDQVGRLTAVARDEGDFLGEQFMQWFLQEQIEEVCLMATLVRVADRAGANLFELENFV
AREVDVAPAASGAPHAAGGRCC (SEQ ID NO: 288)

Ferr_Mt_DS06
MTEYEGPKTKFHALMQEQIHNEFTAAQQYVCIAVYFDSEDLPQLAKHFYSQAVEERNHAMMLVQHLLDRDLRVECPGVDT
VRNQFDRPREALALALDQERTVTDQVGRLCAVARDEGDCLGEQFMQWFLQEQIEEVALMATLVRVADRAGANLFELENFV
AREVDVAPAASGAPHAAGGRL (SEQ ID NO: 289)

Ferr_Mt_DS07
MTEYEGPKTKFHALMQEQIHNEFTAAQQYVAIAVYFDSEDLPQLAKHFYSQAVEERNHAMMLVQHLLDRDLRVEIPGCDT
VRNQFDRPREALALALDQERTVTDQVGRLCAVARDEGDCLGEQFMQWFLQEQIEEVALMATLVRVADRAGANLFELENFV
AREVDVAPAASGAPHAAGGRL (SEQ ID NO: 290)

Ferr_Mt_DS08
MTEYEGPKTKFHALMQEQIHNEFTAAQQYVAIAVYFDSEDLPQLAKHFYSQAVEERNHAMMLVQHLLDRDLRVCIPGVDT
VRCQFDRPREALALALDQERTVTDQVGRLCAVARDEGDCLGEQFMQWFLQEQIEEVALMATLVRVADRAGANLFELENFV
AREVDVAPAASGAPHAAGGRL (SEQ ID NO: 291)

Ferr_Mt_DS09
MTEYEGPKTKFHALMQEQIHNEFTAAQQYVAIAVYFDSEDLPQLAKHFYSQAVEERNHAMMLVQHLLDRDLRVEIPGVDT
VRNQFDRPREALALALDQERTVTDQVGRLCAVARDEGDCLGEQFMQWFLQEQIEEVALMATLVRVADRAGANLFELENFV
AREVDVAPAASGAPHAAGGRC (SEQ ID NO: 292)

Ferr_Mt_DS10
MTEYEGPKTKFHALMQEQIHNEFTAAQQYVAIAVYFDSEDLPQLAKHFYSQAVEERNHAMMLVQHLLDRDLRVEIPGVDT
VRNQFDRPREALALALDQERTVTDQVGRLCAVARDEGDCLGEQFMQWFLQEQIEEVALMATLVRVADRAGANLFELENFV
AREVDVAPAASGAPHAAGGRCC (SEQ ID NO: 293)

Ferr_ec_DS01
MLKPEMIEKLNEQMNLELYSSLLYQQMSAWCCYHTFEGAAAFLRRHAQEEMTHMQRLFDYLCDTGNLPRINTVESPFAEY
SSLDELFQETYKHEQLITQKINELCHAAMTNQDCPTFNFLQWYVSEQHEEEKLFKSIIDKLSLAGKSGEGLYFIDKELST
LDTQN (SEQ ID NO: 294)

Ferr_ec_DS02
MLKPEMIEKLNEQMNLELYSSLLYQQMSAWCSYHTFEGAAAFLRRHAQEEMTHMQRLFDYLTDTGNLPRINCVECPFAEY
SSLDELFQETYKHEQLITQKINELCHAAMTNQDCPTFNFLQWYVSEQHEEEKLFKSIIDKLSLAGKSGEGLYFIDKELST
LDTQN (SEQ ID NO: 295)

Ferr_ec_DS03
MLKPEMIEKLNEQMNLELYSSLLYQQMSAWCSYHTFEGAAAFLRRHAQEEMTHMQRLFDYLTDTGNLPRINTCESPFAEY
SSLDELFQETYKHEQLITQKINELCHAAMTNQDCPTFNFLQWYVSEQHEEEKLFKSIIDKLSLAGKSGEGLYFIDKELST
LDTQN (SEQ ID NO: 296)

Ferr_ec_DS04

MLKPEMIEKLNEQMNLELYSSLLYQQMSAWCCYHTFEGAAAFLRRHAQEEMTHMQRLFDYLCDTGNLPRINTVESPFAEY
SSLDELFQETYKHEQLITQKINELAHAAMTNQDYCTFNFLQWYCSEQHEEEKLFKSIIDKLSLAGKSGEGLYFIDKELST
LDTQN (SEQ ID NO: 297)

Ferr_ec_DS05
MLKPEMIEKLNEQMNLELYSSLLYQQMSAWCSYHTFEGAAAFLRRHAQEEMTHMQRLFDYLTDTGNLPRINCVECPFAEY
SSLDELFQETYKHEQLITQKINELAHAAMTNQDYCTFNFLQWYCSEQHEEEKLFKSIIDKLSLAGKSGEGLYFIDKELST
LDTQN (SEQ ID NO: 298)

Ferr_ec_DS06
MLKPEMIEKLNEQMNLELYSSLLYQQMSAWCSYHTFEGAAAFLRRHAQEEMTHMQRLFDYLTDTGNLPRINTCESPFAEY
SSLDELFQETYKHEQLITQKINELAHAAMTNQDYCTFNFLQWYCSEQHEEEKLFKSIIDKLSLAGKSGEGLYFIDKELST
LDTQN (SEQ ID NO: 299)

Ferr_ec_DS07
MLKPEMIEKLNEQMNLELYSSLLYQQMSAWCCYHTFEGAAAFLRRHAQEEMTHMQRLFDYLCDTGNLPRINTVESPFAEY
SSLDELFQETYKHEQLITQKINELAHAAMTNQDYCTFNFLQWYVSEQCEEEKLFKSIIDKLSLAGKSGEGLYFIDKELST
LDTQN (SEQ ID NO: 300)

Ferr_ec_DS08
MLKPEMIEKLNEQMNLELYSSLLYQQMSAWCCYHTFEGAAAFLRRHAQEEMTHMQRLFDYLCDTGNLPRINTVESPFAEY
SSLDELFQETYKHEQLITQKINELAHAAMTNQDYCTFNFLQWYVSEQCGEEEKLFKSIIDKLSLAGKSGEGLYFIDKELS
TLDTQN (SEQ ID NO: 301)

Ferr_ec_DS09
MLKPEMIEKLNEQMNLELYSSLLYQQMSAWCSYHTFEGAAAFLRRHAQEEMTHMQRLFDYLTDTGNLPRINCVECPFAEY
SSLDELFQETYKHEQLITQKINELAHAAMTNQDYCTFNFLQWYVSEQCEEEKLFKSIIDKLSLAGKSGEGLYFIDKELST
LDTQN (SEQ ID NO: 302)

Ferr_ec_DS10
MLKPEMIEKLNEQMNLELYSSLLYQQMSAWCSYHTFEGAAAFLRRHAQEEMTHMQRLFDYLTDTGNLPRINCVECPFAEY
SSLDELFQETYKHEQLITQKINELAHAAMTNQDYCTFNFLQWYVSEQCGEEEKLFKSIIDKLSLAGKSGEGLYFIDKELS
TLDTQN (SEQ ID NO: 303)

Ferr_ec_DS11
MLKPEMIEKLNEQMNLELYSSLLYQQMSAWCSYHTFEGAAAFLRRHAQEEMTHMQRLFDYLTDTGNLPRINTCESPFAEY
SSLDELFQETYKHEQLITQKINELAHAAMTNQDYCTFNFLQWYVSEQCGEEEKLFKSIIDKLSLAGKSGEGLYFIDKELS
TLDTQN (SEQ ID NO: 304)

Ferr_frog_DS01
MVSQCRQNYHSDCEAAVNRMLNLELYASYTYSSMYCFFDRDDVALHNVAEFFKEHSHEEREHAEKFMKYQNKRGGRCVLQ
DIKKPERDEWGNTLEAMQAALQLEKTVNQALLDLHKLATDKVDPHLCDFLESEYLEEQVKDIKRICDFITNLKRLGLPEN
GMGEYLFDKHSVKESS (SEQ ID NO: 305)

In some embodiments, the fusion proteins of the self-assembling protein nanoparticle carrier comprise any of the disclosed heterologous carrier proteins fused to a ferritin subunit including an amino acid sequence at least 80% (such as at least 85%, at least 90%, at least 95%, or at least 97%) identical to amino acid sequence set forth as any one of SEQ ID NOs: 14-20 or 258-305.

In additional embodiments, any of the disclosed heterologous carrier proteins can be linked to an insect ferritin subunit to construct the self-assembling ferritin nanoparticle carrier including the ferritin nanoparticle fused to the plurality of the heterologous carrier proteins. Insect ferritin protein nanopartciles and their use and production are described, for example, in PCT. Pub. No. WO 2018/005558, which is incorporated by reference herein. Unlike bacterial ferritin, insect ferritin includes twelve copies of two different subunits (termed heavy and light chains; 24 subunits total). The insect ferritin heavy chains trimerize and the insect ferritin light chains trimerize (forming four trimers of heavy chains and four trimers

of light chains) and self-assemble into the globular nanoparticle. In several embodiments, each insect ferritin heavy chain includes an N-terminal fusion to a first heterologous carrier protein, and each insect ferritin light chain includes an N-terminal fusion to a second heterologous carrier protein. This allows for display of two diverse carrier proteins on the same ferritin nanoparticle.

In several embodiments, the insect ferritin heavy and light chains can be from the *Lepidoptera* order of insects, such as ferritin heavy and light chains from Trichoplusia (such as *Trichoplusia ni*), or ferritin heavy and light chains from manduca. Exemplary ferritin heavy and light chain amino acid sequences for *Trichoplusia ni* and manduca proteins are provided below:

Exemplary insect ferritin heavy and light chain sequences with N-terminal truncations that can be included in the fusion protein are set forth below:

Trichoplusia ni ferritin heavy chain with 18-aa N-terminal truncation (nt19)
RSCRNSMRQQIQMEVGASLQYLAMGAHFSKDVVNRPGFAQLFFDAASEEREHAMKLIEYLLMRGELTNDVSSLLQVRPPT
RSSWKGGVEALEHALSMESDVTKSIRNVIKACEDDSEFNDYHLVDYLTGDFLEEQYKGQRDLAGKASTLKKLMDRHEALG
EFIFDKKLLGIDV (SEQ ID NO: 21)

Trichoplusia ni ferritin light chain with 29-aa N-terminal truncation (nt30)
EYGSHGNVATELQAYAKLHLERSYDYLLSAAYFNNYQTNRAGFSKLFKKLSDEAWSKTIDIIKHVTKRGDKMNFDQHSTM
KTERKNYTAENHELEALAKALDTQKELAERAFYIHREATRNSQHLHDPEIAQYLEEEFIEDHAEKIRTLAGHTSDLKKFI
TANNGHDLSLALYVFDEYLQKTV (SEQ ID NO: 22)

Manduca ferritin heavy chain with 38-aa truncation (nt39)
RSCRDSMRRQIQMEVGASLQYLAMGAHFSKDKINRPGFAKLFFDAAGEEREHAMKLIEYLLMRGELTNDVTSLIQVRAPQ
RNKWEGGVDALEHALKMESDVTKSIRTVIKACEDDPEFNDYHLVDYLTGEFLEEQYKGQRDLAGKASTLKKMLDRNSALG
EFIFDKKLMGMDI (SEQ ID NO: 23)

Manduca ferritin light chain with 48-aa N-terminal truncation (nt49)
EYGHHGNVAKEMQAYAALHLERSYEYLLSSSYFNNYQTNRAGFSKLFRKLSDDAWEKTIDLIKHITMRGDEMNFAQRSTQ
KSVDRKNYTVELHELESLAKALDTQKELAERAFFIHREATRNSQHLHDPEVAQYLEEEFIEDHAKTIRNLAGHTTDLKRF
VSGDNGQDLSLALYVFDEYLQKTV (SEQ ID NO: 24)

In some embodiments, the insect ferritin heavy chain can be a *Trichoplusia ni* ferritin heavy chain with an 18 amino acid N-terminal truncation and the insect ferritin light chain can be a *Trichoplusia ni* ferritin light chain with a 29 amino acid N-terminal truncation. For example, the insect ferritin heavy chain comprises an amino acid sequence at least 90% identical to SEQ ID NO: 21, and the insect ferritin light chain comprises an amino acid sequence at least 90% identical SEQ ID NO: 22 In some embodiments, the insect ferritin heavy chain comprises an amino acid sequence set forth as SEQ ID NO: 21, and the insect ferritin light chain comprises an amino acid sequence set forth as SEQ ID NO: 22.

In some embodiments, the insect ferritin heavy chain can be a *manduca* ferritin heavy chain with a 38 amino acid N-terminal truncation and the insect ferritin light chain can be a *manduca* ferritin light chain with a 48 amino acid N-terminal truncation. For example, the insect ferritin heavy chain comprises an amino acid sequence at least 90% identical to SEQ ID NO: 23, and the insect ferritin light chain comprises an amino acid sequence at least 90% identical to SEQ ID NO: 24. In some embodiments, the insect ferritin heavy chain comprises an amino acid sequence set forth as SEQ ID NO: 23, and the insect ferritin light chain comprises an amino acid sequence set forth as SEQ ID NO: 24.

b. Lumazine Synthase (LS)

In some embodiments, any of the disclosed heterologous carrier proteins (such as an rTT, CRM197, or HiD carrier protein) can be linked to a lumazine synthase subunit to construct a self-assembling lumazine synthase nanoparticle carrier including a lumazine synthase nanoparticle fused to a plurality of the heterologous carrier proteins. Lumazine synthase nanoparticles are formed from 60 copies of the lumazine synthase subunit.

The globular form of lumazine synthase nanoparticle is made up of monomeric subunits; non-limiting examples of the sequence of lumazine synthase subunits are provided as:

MQIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAAGELARKEDIDAVIAIGV LIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTLEQAIERAGTKHGNKGWEAALSAIEMANLFKSLR (SEQ ID NO: 25)

QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGCIDCIVRHGGREEDITLVRVPGSWEIPVAAGELARKEDIDAVIAIGVL IRGATPHFDYIASEVSKGLANLSLELRKPITFGVITADTLEQAIERAGTKHGNKCWEAALSAIEMANLFKSLR (SEQ ID NO: 26)

QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAAGELARKENISAVIAIGVL IRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTLEQAIERAGTKHGNKGWEAALSAIEMANLFKSLR (SEQ ID NO: 27)

QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDCIVRHGGREEDITLVRVPGSWEIPVAAGELARKEDIDAVIAIGVL IRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTLEQAIERAGTKHGNKGWEAALSAIEMANLFKSLR (SEQ ID NO: 28)

Additional lumazine synthase subunits are provided with one or more cysteine substitutions to introduce non-native disulfide bond(s) that stabilize the lumazine synthase nanoparticle formed from self-assembled subunits. In some embodiments, the non-native disulfide bond(s) are introduced with L121C-K131C, L121CG-K131C, L121GC-K131C, K7C-R40C, I3C-L50C, I82C-K131CG, E5C-R52C, or E95C-A101C substitutions, or a combination thereof (such as I3C-L50C and I82C-K131CG; E5C-R52C and I82C-K131CG; or E95C-A101C and I82C-K131CG). The residues numbering is with reference to the lumazine synthase subunit set forth as SEQ ID NO: 25. Non-limiting examples include:

LS-L121C-K131C
QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAAGELARKE**nIs**AVIAIGVL IRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADT**c**EQAIERAGT**c**HGNKGWEAALSAIEMANLFKSLR (SEQ ID NO: 306)

LS-L121CG-K131C
QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAAGELARKE**nIs**AVIAIGVL IRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADT**cg**EQAIERAGT**c**HGNKGWEAALSAIEMANLFKSLR (SEQ ID NO: 307)

LS-L121GC-K131C
QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAAGELARKE**nIs**AVIAIGVL IRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADT**gc**EQAIERAGT**c**HGNKGWEAALSAIEMANLFKSLR (SEQ ID NO: 308)

LS-K7C-R40C
QIYEG**c**LTAEGLRFGIVASRFNHALVDRLVEGAIDAIV**c**HGGREEDITLVRVPGSWEIPVAAGELARKE**nIs**AVIAIGVL IRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTLEQAIERAGTKHGNKGWEAALSAIEMANLFKSLR (SEQ ID NO: 309)

LS_Aq_DS01 (I3C-L50C, I82C-K131CG)
QCYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDCIVRHGGREEDITCVRVPGSWEIPVAAGELARKE**DID**AVIAIGVL
CRGATPHFDYIASEVSKGLACLSLELRKPITFGVITADTLEQAIERAGTCGHGNKGWEAALSAIEMANLFKSLR (SEQ ID NO: 310)

LS_Aq_DS02 (E5C-R52C, I82C-K131CG)
QIYCGKLTAEGLRFGIVASRFNHALVDRLVEGAIDCIVRHGGREEDITLVCVPGSWEIPVAAGELARKE**DID**AVIAIGVL
CRGATPHFDYIASEVSKGLACLSLELRKPITFGVITADTLEQAIERAGTCGHGNKGWEAALSAIEMANLFKSLR (SEQ ID NO: 311)

LS_Aq_DS03 (E95C-A101C, I82C-K131CG)
QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDCIVRHGGREEDITLVRVPGSWEIPVAAGELARKE**DID**AVIAIGVL
CRGATPHFDYIASCVSKGLCCLSLELRKPITFGVITADTLEQAIERAGTCGHGNKGWEAALSAIEMANLFKSLR (SEQ ID NO: 312)

In some embodiments, the fusion proteins of the self-assembling protein nanoparticle carrier comprise any of the disclosed heterologous carrier proteins fused to a lumazine synthase subunit including an amino acid sequence at least 80% (such as at least 85%, at least 90%, at least 95%, or at least 97%) identical to amino acid sequence set forth as any one of SEQ ID NOs: 25-28 or 306-312.

c. DNA starvation/stationary phase protection protein (DPS)

In some embodiments, any of the disclosed heterologous carrier proteins (such as an rTT, CRM197, or HiD carrier protein) can be linked to a subunit of a DNA starvation/stationary phase protection protein (DPS) complex, such as a DPS subunit from *Thermosynechococcus elongates*, *Kineococcuc radiotolerans*, or *Nostoc punctiforme*, to construct a self-assembling DPS nanoparticle carrier including a DPS nanoparticle fused to a plurality of the heterologous carrier proteins. Non-limiting examples of the sequence of DPS subunits that can be included in the fusion proteins of the self-assembling protein nanoparticle carrier are provided as:

DNA starvation/stationary phase protection protein (*Thermosynechococcus elongates*)
SATTTLKEQVLTTLKREQANAVVMYLNYKKYHWLTYGPLFRDLHLLFEEQGSEVFAMIDELAERSLMLDGQPVADPADYL
KVATVTPSSGQLTVKQMIEEAIANHELIITEMHQDAEIATEAGDIGTADLYTRLVQTHQKHRWFLKEFLAKGDGLVS
(SEQ ID NO: 29)

DNA starvation/stationary phase protection protein (*Kineococcuc radiotolerans*)
TTIHDVQTTGLTQDAVTGFDASSRLNAGLQEVLVDLTALHLQGKQAHWNIVGENWRDLHLQLDTLVEAARGFSDDVAERM
RAVGGVPDARPQTVAASRIGDVGPDEIDTRACVEAIVALVRHTVDTIRRVHDPIDAEDPASADLLHAITLELEKQAWMIG
SENRSPRR (SEQ ID NO: 30)

DNA starvation/stationary phase protection protein (*Nostoc punctiforme*)
SETQTLLRNFGNVYDNPVLLDRSVTAPVTEGFNVVLASFQALYLQYQKHHFVVEGSEFYSLHEFFNEAYNQVQDHIHEIG
ERLDGLGGVPVATFSKLAELTCFEQESEGVYSSRQMVENDLAAEQAIIGVIRRQAAQAESLGDRGTRYLYEKILLKTEER
AYHLSHFLAKDSLTLGFVQAAQS (SEQ ID NO: 31)

In some embodiments, the fusion proteins of the self-assembling protein nanoparticle carrier comprise any of the disclosed heterologous carrier proteins fused to a DPS subunit including an amino acid sequence at least 80% (such as at least 85%, at least 90%, at least 95%, or at least 97%) identical to amino acid sequence set forth as any one of SEQ ID NOs: 29-31.

d. Bacteriophage Q Beta Capsid protein (qbeta)

In some embodiments, any of the disclosed heterologous carrier proteins (such as an rTT, CRM197, or HiD carrier protein) can be linked to a subunit of a Bacteriophage Q Beta Capsid protein (qbeta) complex to construct a self-assembling qbeta nanoparticle carrier including a qbeta nanoparticle fused to a plurality of the heterologous carrier proteins. A non-limiting example of the sequence of a qbeta subunit that can be included in the fusion proteins of the self-assembling protein nanoparticle carrier is provided as:

AKLETVTLGNIGKDGKQTLVLNPRGVNPTNGVASLSQAGAVPALEKRVTVSVSQPSRNRKNYKVQVKIQNPTACTANGACDPSVTRQAYADVTFSFTQYSTDEERAFVRTELAALLASPLLIDAIDQLNPAY (SEQ ID NO: 32)

In some embodiments, the fusion proteins of the self-assembling protein nanoparticle carrier comprise any of the disclosed heterologous carrier proteins fused to a qbeta subunit including an amino acid sequence at least 80% (such as at least 85%, at least 90%, at least 95%, or at least 97%) identical to the amino acid sequence set forth as SEQ ID NO: 32.

e. Dihydrolipoyl transacetylase protein (e2p)

In some embodiments, any of the disclosed heterologous carrier proteins (such as an rTT, CRM197, or HiD carrier protein) can be linked to a subunit of a dihydrolipoyl transacetylase protein (e2p) complex to construct a self-assembling e2p nanoparticle carrier including an e2p nanoparticle fused to a plurality of the heterologous carrier proteins. E2p nanoparticles are formed from 60 copies of the e2p subunit; structural information is deposited at the Protein Data Bank No. 1B5S. In the globular e2p nanoparticle, the N-terminus of the subunit is surface exposed and the C-terminus of the subunit is inside the globular nanoparticle. A non-limiting example of the sequence of an ep2 subunit that can be included in the fusion proteins of the self-assembling protein nanoparticle carrier is provided as:

AAAKPATTEGEFPETREKMSGIRRAIAKAMVHSKHTAPHVTLMDEADVTKLVAHRKKFKAIAAEKGIKLTFLPYVVKALVSALREYPVLNTAIDDETEEIIQKHYYNIGIAADTDRGLLVPVIKHADRKPIFALAQEINELAEKARDGKLTPGEMKGASCTITNIGSAGGQWFTPVINHPEVAILGIGRIAEKPIVRDGEIVAAPMLALSLSFDHRMIDGATAQKALNHIKRLLSDPELLM (SEQ ID NO: 33)

In some embodiments, the fusion proteins of the self-assembling protein nanoparticle carrier comprise any of the disclosed heterologous carrier proteins fused to an e2p subunit including an amino acid sequence at least 80% (such as at least 85%, at least 90%, at least 95%, or at least 97%) identical to the amino acid sequence set forth as SEQ ID NO: 33.

f. Phosphopantetheine Adenylyltransferase (6ccq)

In some embodiments, any of the disclosed heterologous carrier proteins (such as an rTT, CRM197, or HiD carrier protein) can be linked to a subunit of a Phosphopantetheine Adenylyltransferase (6ccq) complex to construct a self-assembling 6ccq nanoparticle carrier including a 6ccq nanoparticle fused to a plurality of the heterologous carrier proteins. Phosphopantetheine Adenylyltransferase nanoparticles are

formed from 6 copies of the Phosphopantetheine Adenylyltransferase subunit; structural information is deposited at the Protein Data Bank No. 6CCQ. A non-limiting example of the sequence of a 6ccq subunit that can be included in the fusion proteins of the self-assembling protein nanoparticle carrier is provided as:

MQKRAIYPGTFDPITNGHIDIVTRATQMFDHVILAIAASPSKKPMFTLEERVALAQQATAHLGNVEVVGFSDLMANFARN
QHATVLIRGLRAVADFEYEMQLAHMNRHLMPELESVFLMPSKEWSFISSSLVKEVARHQGDVTHFLPENVHQALMAKLAV
D (SEQ ID NO: 34)

In some embodiments, the fusion proteins of the self-assembling protein nanoparticle carrier comprise any of the disclosed heterologous carrier proteins fused to a 6ccq subunit including an amino acid sequence at least 80% (such as at least 85%, at least 90%, at least 95%, or at least 97%) identical to the amino acid sequence set forth as SEQ ID NO: 34.

g. Glutamate Synthase (1f52)

In some embodiments, any of the disclosed heterologous carrier proteins (such as an rTT, CRM197, or HiD carrier protein) can be linked to a subunit of a Glutamate Synthase (1f52) protein complex to construct a self-assembling Glutamate Synthase nanoparticle carrier including a Glutamate Synthase nanoparticle fused to a plurality of the heterologous carrier proteins. A non-limiting example of the sequence of a Glutamate Synthase subunit that can be included in the fusion proteins of the self-assembling protein nanoparticle carrier is provided as:

EHVLTMLNEHEVKFVDLRFTDTKGKEQHVTIPAHQVNAEFFEEGKMFDGSSIGGWKGINESDMVLMPDASTAVIDPFFAD
STLIIRCDILEPGTLQGYDRDPRSIAKRAEDYLRATGIADTVLFGPEPEFFLFDDIRFGASISGSHVAIDDIEGAWNSST
KYEGGNKGHRPGVKGGYFPVPPVDSAQDIRSEMCLVMEQMGLVVEAHHHEVATAGQNEVATRFNTMTKKADEIQIYKYVV
HNVAHRFGKTATFMPKPMFGDNGSGMHCHMSLAKNGTNLFSGDKYAGLSEQALYYIGGVIKHAKAINALANPTTNSYKRL
VPGYEAPVMLAYSARNRSASIRIPVVASPKARRIEVRFPDPAANPYLCFAALLMAGLDGIKNKIHPGEPMDKNLYDLPPE
EAKEIPQVAGSLEEALNALDLDREFLKAGGVFTDEAIDAYIALRREEDDRVRMTPHPVEFELYYSV (SEQ ID NO:
35)

In some embodiments, the fusion proteins of the self-assembling protein nanoparticle carrier comprise any of the disclosed heterologous carrier proteins fused to a Glutamate Synthase subunit including an amino acid sequence at least 80% (such as at least 85%, at least 90%, at least 95%, or at least 97%) identical to the amino acid sequence set forth as SEQ ID NO: 35.

h. Calcium/calmodulin dependent protein kinase IIa (CaMKIIa), C-terminal fragment (5U6Y)

In some embodiments, any of the disclosed heterologous carrier proteins (such as an rTT, CRM197, or HiD carrier protein) can be linked to a c-terminal fragment of a Calcium/calmodulin dependent protein kinase IIa (CaMKIIa) protein to construct a self-assembling CaMKIIa nanoparticle carrier including a nanoparticle based on the C-terminal fragment of CaMKIIa fused to a plurality of the heterologous carrier proteins. The CaMKIIa nanoparticle is formed from 12 copies of the c-terminal fragment of CaMKIIa subunit; structural information is deposited at the Protein Data Bank No. 5U6Y. The N-terminus of the c-terminal fragment is surface exposed in the globular nanoparticle. Non-limiting examples of CaMKIIa

sequences that can be included in the fusion proteins of the self-assembling protein nanoparticle carrier are provided as:

GGKSGGNKKSDGVKESSESTNTAIEDEDTKVRKQEIIKVTEQLIEAISNGDFESYTKMCDPGMTAFEPEALGNLVEGLDF
HRFYFENLWSRNSKPVHTTILNPHIHLMGDESACIAYIRITQYLDAGGIPRTAQSEETRVWHRRDGKWQIVHFHRSGA
(SEQ ID NO: 36)

GVKESSESTNTAIEDEDTKVRKQEIIKVTEQLIEAISNGDFESYTKMCDPGMTAFEPEALGNLVEGLDFHRFYFENLWSR
NSKPVHTTILNPHIHLMGDESACIAYIRITQYLDAGGIPRTAQSEETRVWHRRDGKWQIVHFHRSGA (SEQ ID NO:
37)

STNTAIEDEDTKVRKQEIIKVTEQLIEAISNGDFESYTKMCDPGMTAFEPEALGNLVEGLDFHRFYFENLWSRNSKPVHT
TILNPHIHLMGDESACIAYIRITQYLDAGGIPRTAQSEETRVWHRRDGKWQIVHFHRSGA (SEQ ID NO: 38)

In some embodiments, the fusion proteins of the self-assembling protein nanoparticle carrier comprise any of the disclosed heterologous carrier proteins fused to a Glutamate Synthase subunit including an amino acid sequence at least 80% (such as at least 85%, at least 90%, at least 95%, or at least 97%) identical to the amino acid sequence set forth as any one of SEQ ID NOs: 36-38.

i. HIV capsid oligomerization domain (HIV)

In some embodiments, any of the disclosed heterologous carrier proteins (such as an rTT, CRM197, or HiD carrier protein) can be linked to a HIV capsid oligomoization domain (HIV) to construct a self-assembling HIV capsid oligomerization domain nanoparticle carrier including a nanoparticle based on the HIV capsid oligomerization domain fused to a plurality of the heterologous carrier proteins. Non-limiting examples of HIV capsid oligomerization domain sequences that can be included in the fusion proteins of the self-assembling protein nanoparticle carrier are provided as:

PIVQNLQGQMVHQAISCLCLNAWVKVVEEKAFSPEVIPMFSALSEGATPQDLNTMLNTVGGHQAAMQMLKETINEEAAEW
DRLHPVHAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMTHNPPIPVGEIYKRWIILGLNKIVRMYSPTSILDIRQGPKEP
FRDYVDRFYKTLRAEQASQEVKNAATETLLVQNANPDCKTILKALGPGATLEEMMTACQGVGGPGHKARV (SEQ ID
NO: 39)

PIVQNLQGQMVHQAISCLCLNAWVKVVEEKAFSPEVIPMFSALSEGATPQDLNTMLNTVGGHQAAMQMLKETINEEAAEW
DRLHPVHAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMTHNPPIPVGEIYKRWIILGLNKIVRMYSPTSILDIRQGPKEP
FRDYVDRFYKTLRAEQASQEVKNAATETLLVQNANPDCKTILKALGPGATLEEMMTA (SEQ ID NO: 40)

In some embodiments, the fusion proteins of the self-assembling protein nanoparticle carrier comprise any of the disclosed heterologous carrier proteins fused to a HIV capsid oligomerization domain including an amino acid sequence at least 80% (such as at least 85%, at least 90%, at least 95%, or at least 97%) identical to the amino acid sequence set forth as any one of SEQ ID NOs: 39-40.

j. Hexamer

In some embodiments, any of the disclosed heterologous carrier proteins (such as an rTT, CRM197, or HiD carrier protein) can be linked to a Hexamer subunit to construct a hexamer nanoparticle carrier including a nanoparticle based on the hexamer sequence fused to a plurality of the heterologous carrier

proteins. A non-limiting examples of a hexamer sequence that can be included in the fusion proteins of the self-assembling protein nanoparticle carrier is provided as:

PTLYNVSLVMSDTAGTCY (SEQ ID NO: 41)

In some embodiments, the fusion proteins of the self-assembling protein nanoparticle carrier comprise any of the disclosed heterologous carrier proteins fused to a hexamer subunit including an amino acid sequence at least 80% (such as at least 85%, at least 90%, at least 95%, or at least 97%) identical to the amino acid sequence set forth as SEQ ID NO: 41.

k. T4 fibritin Foldon domain (Fd)

In some embodiments, any of the disclosed heterologous carrier proteins (such as an rTT, CRM197, or HiD carrier protein) can be linked to a T4 fibritin Foldon domain to construct a hexamer nanoparticle carrier including a nanoparticle based on the T4 fibritin Foldon domain sequence fused to a plurality of the heterologous carrier proteins. A non-limiting examples of a T4 fibritin Foldon domain sequence that can be included in the fusion proteins of the self-assembling protein nanoparticle carrier is provided as:

GYIPEAPRDGQAYVRKDGEWVLLSTFL (SEQ ID NO: 42)

In some embodiments, the fusion proteins of the self-assembling protein nanoparticle carrier comprise any of the disclosed heterologous carrier proteins fused to a T4 fibritin Foldon domain including an amino acid sequence at least 80% (such as at least 85%, at least 90%, at least 95%, or at least 97%) identical to the amino acid sequence set forth as SEQ ID NO: 42.

l. Encapsulin

In some embodiments, any of the disclosed heterologous carrier proteins (such as an rTT, CRM197, or HiD carrier protein) can be linked to an encapsulin subunit to construct a self-assembling encapsulin nanoparticle carrier including an encapsulin nanoparticle fused to a plurality of the heterologous carrier proteins. Encapsulin nanoparticles are formed from 60 copies of the encapsulin subunit.

The globular form of the encapsulin nanoparticle is made up of monomeric subunits. A non-limiting example of the sequence of an encapsulin subunit is provided as:

MEFLKRSFAPLTEKQWQEIDNRAREIFKTQLYGRKFVDVEGPYGWEYAAHPLGEVEVLSDENEVVKWGLRKSLPLIELRA
TFTLDLWELDNLERGKPNVDLSSLEETVRKVAEFEDEVIFRGCEKSGVKGLLSFEERKIECGSTPKDLLEAIVRALSIFS
KDGIEGPYTLVINTDRWINFLKEEAGHYPLEKRVEECLRGGKIITTPRIEDALVVSERGGDFKLILGQDLSIGYEDREKD
AVRLFITETFTFQVVNPEALILLKF (SEQ ID NO: 43)

Additional encapsulin subunits are provided with one or more cysteine substitutions to introduce non-native disulfide bond(s) that stabilize the encapsulin nanoparticle formed from self-assembled subunits. In some embodiments, the non-native disulfide bond(s) are introduced with G53C-R94C, G53C-K96C, or K146C-A185C substitutions, or a combination thereof. The residues numbering is with reference to the encapsulin subunit set forth as SEQ ID NO: 43. Non-limiting examples include:

EN G53C-R94C
MEFLKRSFAPLTEKQWQEIDNRAREIFKTQLYGRKFVDVEGPYGWEYAAHPLCEVEVLSDENEVVKWGLRKSLPLIELRA
TFTLDLWELDNLECGKPNVDLSSLEETVRKVAEFEDEVIFRGCEKSGVKGLLSFEERKIECGSTPKDLLEAIVRALSIFS
KDGIEGPYTLVINTDRWINFLKEEAGHYPLEKRVEECLRGGKIITTPRIEDALVVSERGGDFKLILGQDLSIGYEDREKD
AVRLFITETFTFQVVNPEALILLKF (SEQ ID NO: 313)

EN G53C-K96C
MEFLKRSFAPLTEKQWQEIDNRAREIFKTQLYGRKFVDVEGPYGWEYAAHPLCEVEVLSDENEVVKWGLRKSLPLIELRA
TFTLDLWELDNLErGcPNVDLSSLEETVRKVAEFEDEVIFRGCEKSGVKGLLSFEERKIECGSTPKDLLEAIVRALSIFS
KDGIEGPYTLVINTDRWINFLKEEAGHYPLEKRVEECLRGGKIITTPRIEDALVVSERGGDFKLILGQDLSIGYEDREKD
AVRLFITETFTFQVVNPEALILLKF (SEQ ID NO: 314)

EN K146C-A185C
MEFLKRSFAPLTEKQWQEIDNRAREIFKTQLYGRKFVDVEGPYGWEYAAHPLgEVEVLSDENEVVKWGLRKSLPLIELRA
TFTLDLWELDNLErGkPNVDLSSLEETVRKVAEFEDEVIFRGCEKSGVKGLLSFEERKIECGSTPcDLLEAIVRALSIFS
KDGIEGPYTLVINTDRWINFLKEEcGHYPLEKRVEECLRGGKIITTPRIEDALVVSERGGDFKLILGQDLSIGYEDREKD
AVRLFITETFTFQVVNPEALILLKF (SEQ ID NO: 315)

In some embodiments, the fusion proteins of the self-assembling protein nanoparticle carrier comprise any of the disclosed heterologous carrier proteins fused to an encapsulin subunit including an amino acid sequence at least 80% (such as at least 85%, at least 90%, at least 95%, or at least 97%) identical to amino acid sequence set forth as SEQ ID NO: 43 or 313-315.

Encapsulin proteins are a conserved family of bacterial proteins also known as linocin-like proteins that form large protein assemblies that function as a minimal compartment to package enzymes. The encapsulin assembly is made up of monomeric subunits, which are polypeptides having a molecule weight of approximately 30 kDa. Following production, the monomeric subunits self-assemble into the globular encapsulin assembly including 60, or in some cases, 180 monomeric subunits. Methods of constructing encapsulin nanoparticles are described, for example, in Sutter *et al.* (Nature Struct. and Mol. Biol., 15:939-947, 2008, which is incorporated by reference herein in its entirety). In specific examples, the encapsulin polypeptide is bacterial encapsulin, such as *Thermotoga maritime* or *Pyrococcus furiosus* or *Rhodococcus erythropolis* or *Myxococcus xanthus* encapsulin.

m. Acinetobacter phage AP205 (AP205)

In some embodiments, any of the disclosed heterologous carrier proteins (such as an rTT, CRM197, or HiD carrier protein) can be linked to a *Acinetobacter* phage AP205 domain to construct a self-assembing nanoparticle carrier including a nanoparticle based on the *Acinetobacter* phage AP205 domain sequence fused to a plurality of the heterologous carrier proteins. A non-limiting examples of an *Acinetobacter* phage AP205 domain sequence that can be included in the fusion proteins of the self-assembling protein nanoparticle carrier is provided as:

AP205
MANKPMQPITSTANKIVWSDPTRLSTTFSASLLRQRVKVGIAELNNVSGQYVSVYKRPAPKPEGCADACVIMPNENQSIR
TVISGSAENLATLKAEWETHKRNVDTLFASGNAGLGFLDPTAAIVSSDTT (SEQ ID NO: 316)

Additional *Acinetobacter* phage AP205 subunits are provided with one or more cysteine substitutions to introduce non-native disulfide bond(s) that stabilize the *Acinetobacter* phage

AP205nanoparticle formed from self-assembled subunits. In some embodiments, the non-native disulfide bond(s) are introduced with T81C (which forms a disulfide with a cysteine already present in AP205), S53C-H100C, or V82C-R80C substitutions, or a combination thereof. The residues numbering is with reference to the *Acinetobacter* phage AP205 subunit set forth as SEQ ID NO: 316. Non-limiting examples include:

AP205-T81C
MANKPMQPITSTANKIVWSDPTRLSTTFSASLLRQRVKVGIAELNNVSGQYVSVYKRPAPKPEGCADACVIMPNENQSIR
cVISGSAENLATLKAEWETHKRNVDTLFASGNAGLGFLDPTAAIVSSDTT (SEQ ID NO: 317)

AP205 S53C-H100C
MANKPMQPITSTANKIVWSDPTRLSTTFSASLLRQRVKVGIAELNNVSGQYVcVYKRPAPKPEGCADACVIMPNENQSIR
TVISGSAENLATLKAEWETcKRNVDTLFASGNAGLGFLDPTAAIVSSDTT (SEQ ID NO: 318)

AP205 V82C-R80C
MANKPMQPITSTANKIVWSDPTRLSTTFSASLLRQRVKVGIAELNNVSGQYVSVYKRPAPKPEGCADACVIMPNENQSIc
bcISGSAENLATLKAEWETHKRNVDTLFASGNAGLGFLDPTAAIVSSDTT (SEQ ID NO: 319)

AP205 C65-C69GC
MANKPMQPITSTANKIVWSDPTRLSTTFSASLLRQRVKVGIAELNNVSGQYVSVYKRPAPKPEGCADAgCVIMPNENQSI
RTVISGSAENLATLKAEWETHKRNVDTLFASGNAGLGFLDPTAAIVSSDTT (SEQ ID NO: 320)

In some embodiments, the fusion proteins of the self-assembling protein nanoparticle carrier comprise any of the disclosed heterologous carrier proteins fused to a *Acinetobacter* phage AP205 subunit including an amino acid sequence at least 80% (such as at least 85%, at least 90%, at least 95%, or at least 97%) identical to the amino acid sequence set forth as SEQ ID NO: 316-320.

n. Hepatitis B capsid protein (HBV)

In some embodiments, any of the disclosed heterologous carrier proteins (such as an rTT, CRM197, or HiD carrier protein) can be linked to a Hepatitis B capsid protein domain to construct a self-assembling nanoparticle carrier including a nanoparticle based on the Hepatitis B capsid protein domain sequence fused to a plurality of the heterologous carrier proteins. A non-limiting examples of an Hepatitis B capsid protein domain sequence that can be included in the fusion proteins of the self-assembling protein nanoparticle carrier is provided as:

HBV
MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAILCWGELMTLATWVGVNLEDPA
SRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT
PSPRRRRSQSPRRRRSQSRESQC (SEQ ID NO: 321)

Additional Hepatitis B capsid protein subunits are provided with one or more cysteine substitutions to introduce non-native disulfide bond(s) that stabilize the Hepatitis B capsid protein domain nanoparticle formed from self-assembled subunits. In some embodiments, the non-native disulfide bond(s) are introduced with P25C-R127C, E14C-A36C, D29C-R127C, F18C-A36C, or D29C-R127C substitutions, or a combination thereof. The residues numbering is with reference to the Hepatitis B capsid protein subunit set forth as SEQ ID NO: 321. Non-limiting examples include:

HBV P25C-R127C
MDIDPYKEFGATVELLSFLPSDFFcSVRDLLDTASALYREALESPEHCSPHHTALRQAILCWGELMTLATWVGVNLEDPA
SRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETVIEYLVSFGVWIcTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT
PSPRRRRSQSPRRRRSQSRESQC (SEQ ID NO: 322)

HBV E14C-A36C
MDIDPYKEFGATVcLLSFLPSDFFPSVRDLLDTAScLYREALESPEHCSPHHTALRQAILCWGELMTLATWVGVNLEDPA
SRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT
PSPRRRRSQSPRRRRSQSRESQC (SEQ ID NO: 323)

HBV D29C-R127C
MDIDPYKEFGATVELLSFLPSDFFPSVRcLLDTASALYREALESPEHCSPHHTALRQAILCWGELMTLATWVGVNLEDPA
SRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETVIEYLVSFGVWIcTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT
PSPRRRRSQSPRRRRSQSRESQC (SEQ ID NO: 324)

HBV_DS01 (F18C-A36C)
MDIDPYKEFGATVELLScLPSDFFPSVRDLLDTAScLYREALESPEHCSPHHTALRQAILCWGELMTLATWVGVNLEDPA
SRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT
PSPRRRRSQSPRRRRSQSRESQC (SEQ ID NO: 325)

HBV_DS02 (D29C-R127C)
MDIDPYKEFGATVELLSFLPSDFFPSVRcLLDTASALYREALESPEHCSPHHTALRQAILCWGELMTLATWVGVNLEDPA
SRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETVIEYLVSFGVWIcTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT
PSPRRRRSQSPRRRRSQSRESQC (SEQ ID NO: 326)

In some embodiments, the fusion proteins of the self-assembling protein nanoparticle carrier comprise any of the disclosed heterologous carrier proteins fused to a Hepatitis B capsid subunit including an amino acid sequence at least 80% (such as at least 85%, at least 90%, at least 95%, or at least 97%) identical to the amino acid sequence set forth as any one of SEQ ID NO: 321-326.

2. Heterologous Carrier Proteins

The heterologous carrier protein included in the fusion protein can be any carrier protein suitable for use as with a vaccine that is a single polypeptide chain of amino acids (as opposed to a protein complex). Examples of suitable heterologous carrier proteins are those that can increase the immunogenicity of the conjugate and/or elicit antibodies against the carrier which are diagnostically, analytically, and/or therapeutically beneficial. Specific, non-limiting examples of suitable polypeptide carriers include, but are not limited to, natural, semi-synthetic or synthetic polypeptides or proteins from bacteria or viruses. In one embodiment, bacterial products for use as carriers include bacterial toxins, such as those that are single polypeptide chains (or a fragment thereof) that mediate toxic effects, inflammatory responses, stress, shock, chronic sequelae, or mortality in a susceptible host. Specific, non-limiting examples of such bacterial toxins include, but are not limited to: single polypeptide chains of *B. anthracis* PA (for example, as encoded by bases 143779 to 146073 of GENBANK® Accession No. NC 007322); *B. anthracis* LF (for example, as encoded by the complement of bases 149357 to 151786 of GENBANK® Accession No. NC 007322); bacterial toxins and toxoids, such as tetanus toxin/toxoid (for example, as described in U.S. Patent Nos. 5,601,826 and 6,696,065); diphtheria toxin/toxoid (for example, as described in U.S. Patent Nos. 4,709,017 and 6,696,065), such as tetanus toxin heavy chain C fragment; *P. aeruginosa* exotoxin/toxoid (for example, as described in U.S. Patent Nos. 4,428,931, 4,488,991 and 5,602,095); pertussis toxin/toxoid (for example,

as described in U.S. Patent Nos. 4,997,915, 6,399,076 and 6,696,065); and *C. perfringens* exotoxin/toxoid (for example, as described in U.S. Patent Nos. 5,817,317 and 6,403,094) *C. difficile* toxin B or A, or analogs or mimetics of and combinations of two or more thereof. Viral proteins, such as hepatitis B surface antigen (for example, as described in U.S. Patent Nos. 5,151,023 and 6,013,264) and core antigen (for example, as described in U.S. Patent Nos. 4,547,367 and 4,547,368) can also be used as carriers, as well as single polypeptide chains of proteins from higher organisms such as keyhole limpet hemocyanin (KLH), horseshoe crab hemocyanin, Concholepas Concholepas Hemocyanin (CCH), Ovalbumin (OVA), edestin, mammalian serum albumins (such as bovine serum albumin), and mammalian immunoglobulins.

In some embodiments, the heterologous carrier protein is selected from one of: a Keyhole Limpet Hemocyanin (KLH) subunit, recombinant tetanus toxin heavy chain C fragment (rTT), diphtheria toxin variant CRM197, or H influenzae protein D (HiD). CRM197 is a genetically detoxified form of diphtheria toxin; a single mutation at position 52, substituting glutamic acid for glycine, causes the ADP-ribosyltransferase activity of the native diphtheria toxin to be lost. For description of exemplary protein carriers for vaccines, see Pichichero, Protein carriers of conjugate vaccines: characteristics, development, and clinical trials, Hum Vaccin Immunother., 9: 2505-2523,2013, which is incorporated by reference herein in its entirety).

In some embodiments, the heterologous carrier protein is an rTT protein, for example, comprising the amino acid sequence set forth as:

MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAIHLVNNESSEVIVHKAMDIEY
NDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKF
NAYLANKWVFITITNDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEK
LYTSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPN
NEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKNA
SLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND (SEQ ID NO: 44)

In some embodiments, the heterologous carrier protein is an rTT protein comprising amino acid substitutions to remove one or more N-linked glycosylation sites. It is believed that removal of the N-linked glycosylation sites may improve accessibility of the protein surface for conjugation to the HIV-1 Env fusion peptide. Exemplary rTT protein sequences with modifications to remove one or more N-linked glycosylation sites are provided as:

NLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNS**A**VITYPDAQLVPGINGKAIHLVNNE**A**SEVIVHKAMDIEYND
MFN**Q**FTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNA
YLANKWVFITITNDRLSSANLYINGVLMGSAEITGLGAIREDN**Q**ITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLY
TSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLK**Q**ITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNE
IDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKNASL
GLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND (SEQ ID NO: 45)

NLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHLVNNEASEVIVHKAMDIEYND
MFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNA
YLANKWVFITITNDRLSSANLYINGVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLY
TSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNE
IDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKQASL
GLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND (SEQ ID NO: 46)

In some embodiments, the heterologous carrier protein is an rTT protein comprising amino acid substitutions to remove one or more N-linked glycosylation sites as well as to introduce lysine residues at surface exposed positions of the carrier. Increasing the number of lysine residues in the heterologous carrier protein increases the number of available sites for conjugation to the HIV-1 Env fusion peptides with methods targeting the amino moiety of lysine, such as sulfosuccinimidyl (4-iodoacetyl)aminobenzoate (Sulfo-SIAB) linkers. Exemplary rTT protein sequences with modifications to remove one or more N-linked glycosylation sites and/or to add lysine residues are provided as:

NLDCWVDkEEDIDVILKKSTILNLDINkDIISDISkFNSAVITYPDAQLVPGINGKAIHLVNNEkSEVIVHKAMkIEYND
MFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNA
YLANKWVFITITNDRLSSANLYINGVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLY
TSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNE
IDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKQASL
GLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND (SEQ ID NO: 47)

NLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAIHLVNNESSEVIVHKAMDIEYND
MFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNA
YLANKWVFITITNDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLY
TSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNE
IDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKNASL
GLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND (SEQ ID NO: 48)

NLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNS**A**VITYPDAQLVPGINGKAIHLVNNE**A**SEVIVHKAMDIEYND
MFN**Q**FTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNA
YLANKWVFITITNDRLSSANLYINGVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLY
TSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNE
IDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKNASL
GLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND (SEQ ID NO: 49)

NLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHLVNNEASEVIVHKAMDIEYND
MFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNA
YLANKWVFITITNDRLSSANLYINGVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLY
TSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNE
IDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKQASL
GLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND (SEQ ID NO: 50)

NLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHLVNNEASEVIVHKAMDIEYND
MFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNA
YLANKWVFITITNDRLSSANLYINGVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLY
TSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNE
IDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKQASL
GLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND (SEQ ID NO: 51)

NLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHLVNNEASEVIVHKAMDIEYND
MFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNA
YLANKWVFITITNDRLSSANLYINGVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLY
TSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNE
IDSFVKSGDFIKLYVSYkNNEHIVGYPKDGNAFNkLDRILRVGYkAPkIPLYKKMEAVKLRDLKTYSVQLKLYDDKQASL
GLVGTHNGQIGkDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND (SEQ ID NO: 52)

NLDCWVDkEEDIDVILKKSTILNLDINkDIISDISkFNSAVITYPDAQLVPGINGKAIHLVNNEkSEVIVHKAMkIEYND
MFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNA
YLANKWVFITITNDRLSSANLYINGVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLY
TSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNE
IDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKQASL
GLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND (SEQ ID NO: 53)

NLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHLVNNEASEVIVHKAMDIEYND
MFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHkkSIGSGWSVSLKGNNLIWTLKDSkGEVRQITFRDLPkKFNA

YLANKWVFITITNDRkSSANLYINGVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLY
TSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNE
IDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKQASL
GLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND (SEQ ID NO: 54)

NLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHLVNNEASEVIVHKAMDIEYND
MFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNA
YLANKWVFITITNDRLSSANLYINGVLMGSAEITkLGAIREDNQITLKLDRCkNNNQYVSIDKFRIFCKALNPKEIEKLY
TSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTkAPSYTNGKLNIYYRRLYNGLKFIIKRYkPNNk
IDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKQASL
GLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND (SEQ ID NO: 55)

NLDCWVDkEEDIDVILKKSTILNLDINkDIISDISkFNSAVITYPDAQLVPGINGKAIHLVNNEkSEVIVHKAMkIEYND
MFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHkkSIGSGWSVSLKGNNLIWTLKDSkGEVRQITFRDLPkKFNA
YLANKWVFITITNDRkSSANLYINGVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLY
TSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNE
IDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKQASL
GLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND (SEQ ID NO: 56)

NLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHLVNNEASEVIVHKAMDIEYND
MFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNA
YLANKWVFITITNDRLSSANLYINGVLMGSAEITkLGAIREDNQITLKLDRCkNNNQYVSIDKFRIFCKALNPKEIEKLY
TSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTkAPSYTNGKLNIYYRRLYNGLKFIIKRYkPNNk
IDSFVKSGDFIKLYVSYkNNEHIVGYPKDGNAFNkLDRILRVGYkAPkIPLYKKMEAVKLRDLKTYSVQLKLYDDKQASL
GLVGTHNGQIGkDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND (SEQ ID NO: 57)

NLDCWVDkEEDIDVILKKSTILNLDINkDIISDISkFNSAVITYPDAQLVPGINGKAIHLVNNEkSEVIVHKAMkIEYND
MFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNA
YLANKWVFITITNDRLSSANLYINGVLMGSAEITkLGAIREDNQITLKLDRCkNNNQYVSIDKFRIFCKALNPKEIEKLY
TSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTkAPSYTNGKLNIYYRRLYNGLKFIIKRYkPNNk
IDSFVKSGDFIKLYVSYkNNEHIVGYPKDGNAFNkLDRILRVGYkAPkIPLYKKMEAVKLRDLKTYSVQLKLYDDKQASL
GLVGTHNGQIGkDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND (SEQ ID NO: 58)

NLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHLVNNEASEVIVHKAMDIEYND
MFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHkkSIGSGWSVSLKGNNLIWTLKDSkGEVRQITFRDLPkKFNA
YLANKWVFITITNDRkSSANLYINGVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLY
TSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNE
IDSFVKSGDFIKLYVSYkNNEHIVGYPKDGNAFNkLDRILRVGYkAPkIPLYKKMEAVKLRDLKTYSVQLKLYDDKQASL
GLVGTHNGQIGkDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND (SEQ ID NO: 59)

NLDCWVDkEEDIDVILKKSTILNLDINkDIISDISkFNSAVITYPDAQLVPGINGKAIHLVNNEkSEVIVHKAMkIEYND
MFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHkkSIGSGWSVSLKGNNLIWTLKDSkGEVRQITFRDLPkKFNA
YLANKWVFITITNDRkSSANLYINGVLMGSAEITkLGAIREDNQITLKLDRCkNNNQYVSIDKFRIFCKALNPKEIEKLY
TSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTkAPSYTNGKLNIYYRRLYNGLKFIIKRYkPNNk
IDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKQASL
GLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND (SEQ ID NO: 60)

NLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHLVNNEASEVIVHKAMDIEYND
MFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHkkSIGSGWSVSLKGNNLIWTLKDSkGEVRQITFRDLPkKFNA
YLANKWVFITITNDRkSSANLYINGVLMGSAEITkLGAIREDNQITLKLDRCkNNNQYVSIDKFRIFCKALNPKEIEKLY
TSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTkAPSYTNGKLNIYYRRLYNGLKFIIKRYkPNNk
IDSFVKSGDFIKLYVSYkNNEHIVGYPKDGNAFNkLDRILRVGYkAPkIPLYKKMEAVKLRDLKTYSVQLKLYDDKQASL
GLVGTHNGQIGkDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND (SEQ ID NO: 61)

NLDCWVDkEEDIDVILKKSTILNLDINkDIISDISkFNSAVITYPDAQLVPGINGKAIHLVNNEkSEVIVHKAMkIEYND
MFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNA
YLANKWVFITITNDRLSSANLYINGVLMGSAEITkLGAIREDNQITLKLDRCkNNNQYVSIDKFRIFCKALNPKEIEKLY
TSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTkAPSYTNGKLNIYYRRLYNGLKFIIKRYkPNNk
IDSFVKSGDFIKLYVSYkNNEHIVGYPKDGNAFNkLDRILRVGYkAPkIPLYKKMEAVKLRDLKTYSVQLKLYDDKQASL
GLVGTHNGQIGkDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND (SEQ ID NO: 62)

NLDCWVDkEEDIDVILKKSTILNLDINkDIISDISkFNSAVITYPDAQLVPGINGKAIHLVNNEkSEVIVHKAMkIEYND
MFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHkkSIGSGWSVSLKGNNLIWTLKDSkGEVRQITFRDLPkKFNA
YLANKWVFITITNDRkSSANLYINGVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLY
TSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNE
IDSFVKSGDFIKLYVSYkNNEHIVGYPKDGNAFNkLDRILRVGYkAPkIPLYKKMEAVKLRDLKTYSVQLKLYDDKQASL
GLVGTHNGQIGkDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND (SEQ ID NO: 63)

In some embodiments, the heterologous carrier protein is a fragment of the rTT protein, such as a fragment of rTT protein comprising, consisting essentially of, or consisting of the amino acid sequence set forth as:

NLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAIHLVNNESSEVIVHKAMDIEYND
MFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNA
YLANKWVFITITNDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLY
TSYLSIT (SEQ ID NO: 64)

In some embodiments, the fusion protein can include an rTT sequence set forth as any one of SEQ ID NOs: 44-64, or an amino acid sequence at least 90% identical thereto.

In some embodiments, the heterologous carrier protein is a HiD protein, for example, comprising the amino acid sequence set forth as:

SNMANTQMKSDKIIIAHRGASGYLPEHTLESKALAFAQQADYLEQDLAMTKDGRLVVIHDHFLDGLTDVAKKFPHRHRKD
GRYYVIDFTLKEIQSLEMTENFETKDGKQAQVYPNRFPLWKSHFRIHTFEDEIEFIQGLEKSTGKKVGIYPEIKAPWFHH
QNGKDIAAETLKVLKKYGYDKKTDMVYLQTFDFNELKRIKTELLPQMGMDLKLVQLIAYTDWKETQEKDPKGYWVNYNYD
WMFKPGAMAEVVKYADGVGPGWYMLVNKEESKPDNIVYTPLVKELAQYNVEVHPYTVRKDALPEFFTDVNQMYDALLNKS
GATGVFTDFPDTGVEFLKGIK (SEQ ID NO: 65)

In some embodiments, the fusion protein can include a HiD sequence set forth as SEQ ID NO: 65, or an amino acid sequence at least 90% identical thereto.

In some embodiments, the heterologous carrier protein is a HiD protein comprising amino acid substitutions to remove one or more N-linked glycosylation sites and/or to introduce lysine residues at surface exposed positions of the carrier. It is believed that removal of the N-linked glycosylation sites may improve accessibility of the protein surface for conjugation to the HIV-1 Env fusion peptide.

In some embodiments, the heterologous carrier protein is a CRM197 protein, for example, comprising the amino acid sequence set forth as:

GADDVVDSSKSFVMENFSSYHGTKPGYVDSIQKGIQKPKSGTQGNYDDDWKEFYSTDNKYDAAGYSVDNENPLSGKAGGV
VKVTYPGLTKVLALKVDNAETIKKELGLSLTEPLMEQVGTEEFIKRFGDGASRVVLSLPFAEGSSSVEYINNWEQAKALS
VELEINFETRGKRGQDAMYEYMAQACAGNRVRRSVGSSLSCINLDWDVIRDKTKTKIESLKEHGPIKNKMSESPNKTVSE
EKAKQYLEEFHQTALEHPELSELKTVTGTNPVFAGANYAAWAVNVAQVIDSETADNLEKTTAALSILPGIGSVMGIADGA
VHHNTEEIVAQSIALSSLMVAQAIPLVGELVDIGFAAYNFVESIINLFQVVHNSYNRPAYSPGHKTQPFLHDGYAVSWNT
VEDSIIRTGFQGESGHDIKITAENTPLPIAGVLLPTIPGKLDVNKSKTHISVNGRKIRMRCRAIDGDVTFCRPKSPVYVG
NGVHANLHVAFHRSSSEKIHSNEISSDSIGVLGYQKTVDHTKVNSKLSLFFEIKS (SEQ ID NO: 66)

In some embodiments, the fusion protein can include a CRM197 sequence set forth as SEQ ID NO: 66, or an amino acid sequence at least 90% identical thereto.

In some embodiments, the heterologous carrier protein is a CRM197 protein comprising amino acid substitutions to remove one or more N-linked glycosylation sites. It is believed that removal of the N-linked glycosylation sites may improve accessibility of the protein surface for conjugation to the HIV-1 Env fusion peptide. An exemplary CRM197 protein sequence with modifications to remove one or more N-linked glycosylation sites is provided as:

GADDVVDSSKSFVMENFASYHGTKPGYVDSIQKGIQKPKSGTQGNYDDDWKEFYSTDNKYDAAGYSVDNENPLSGKAGGV
VKVTYPGLTKVLALKVDNAETIKKELGLSLTEPLMEQVGTEEFIKRFGDGASRVVLSLPFAEGSSSVEYINNWEQAKALS
VELEINFETRGKRGQDAMYEYMAQACAGNRVRRSVGSSLSCINLDWDVIRDKTKTKIESLKEHGPIKNKMSESPNKAVSE
EKAKQYLEEFHQTALEHPELSELKTVTGTNPVFAGANYAAWAVNVAQVIDSETADNLEKTTAALSILPGIGSVMGIADGA
VHHNTEEIVAQSIALSSLMVAQAIPLVGELVDIGFAAYNFVESIINLFQVVHNSYNRPAYSPGHKTQPFLHDGYAVSWNT
VEDSIIRTGFQGESGHDIKITAENTPLPIAGVLLPTIPGKLDVNKAKTHISVNGRKIRMRCRAIDGDVTFCRPKSPVYVG
NGVHANLHVAFHRSSSEKIHSNEISSDSIGVLGYQKTVDHTKVNSKLSLFFEIKS (SEQ ID NO: 67)

In some embodiments, the fusion protein can include a CRM197 sequence set forth as SEQ ID NOs: 67, or an amino acid sequence at least 90% identical thereto.

In some embodiments, the heterologous carrier protein is a Meningococcal outer membrane protein complex (OMPC) protein. An exemplary OMPC protein sequence with modifications to remove one or more N-linked glycosylation sites is provided as:

DFTIQDIRVEGLQRTEPSTVFNYLPVKVGDTYNDTHGSAIIKSLYATGFFDDVRVETADGQLLLTVIERPTIGSLNITGA
KMLQNDAIKKNLESFGLAQSQYFNQATLNQAVAGLKEEYLGRGKLNIQITPKVTKLARNRVDIDITIDEGKSAKITDIEF
EGNQVYSDRKLMRQMSLTEGGIWTWLTRSNQFNEQKFAQDMEKVTDFYQNNGYFDFRILDTDIQTNEDKTKQTIKITVHE
GERFRWGKVSIEGDTNEVPKAELEKLLTMKPGKWYERQQMTAVLGEIQNRMGSAGYAYSEISVQPLPNAETKTVDFVLHI
EPGRKIYVNEIHITGNNKTRDEVVRRELRQMESAPYDTSKLQRSKERVELLGYFDNVQFDAVPLAGTPDKVDLNMSLTER
STGSLDLSAGWVQDTGLVMSAGVSQDNLFGTGKSAALRASRSKTTLNGSLSFTDPYFTADGVSLGYDVYGKAFDPRKAST
SIKQYKTTTAGAGIRMSVPVTEYDRVNFGLVAEHLTVNTYNKAPKHYADFIKKYGKTDGTDGSFKGWLYKGTVGWGRNKT
DSALWPTRGYLTGVNAEIALPGSKLQYYSATHNQTWFFPLSKTFTLMFGGEVGIAGGYGKTKEIPFFENFYGGGLGSVRG
YESGTLGPKVYDEYGEKISYGGNKKANVSAELLFPMPGAKDARTVRLSLFADAGSVWDGKTYDDNSSSATGGRVQNIYGA
GNTHKSTFTNELRYSAGGAVTWLSPLGPMKFSYAYPLKKKPEDEIQRFQFQLGTTF (SEQ ID NO: 222)

In some embodiments, the heterologous carrier protein is an OMPC protein comprising amino acid substitutions to remove one or more N-linked glycosylation sites and/or to introduce lysine residues at surface exposed positions of the carrier. It is believed that removal of the N-linked glycosylation sites may improve accessibility of the protein surface for conjugation to the HIV-1 Env fusion peptide.

In some embodiments, the heterologous carrier protein is an Outer-membrane lipoprotein carrier protein comprising amino acid substitutions to remove one or more N-linked glycosylation sites. It is believed that removal of the N-linked glycosylation sites may improve accessibility of the protein surface for conjugation to the HIV-1 Env fusion peptide. An exemplary Outer-membrane lipoprotein carrier protein sequence with modifications to remove one or more N-linked glycosylation sites is provided as:

QAGAVDALKQFNNDADGISGSFTQTVQSKKKTQTAHGTFKILRPGLFKWEYTSPYKQTIVGDGQTVWLYDVDLAQVTKSS
QDQAIGGSPAAILSNKTALESSYTLKEDGSSNGIDYVLATPKRNNAGYQYIRIGFKGGNLAAMQLKDSFGNQTSISFGGL
NTNPQLSRGAFKFTPPKGVDVLSN (SEQ ID NO: 223)

In some embodiments, the heterologous carrier protein is a Outer-membrane lipoprotein carrier protein comprising amino acid substitutions to remove one or more N-linked glycosylation sites and/or to introduce lysine residues at surface exposed positions of the carrier. It is believed that removal of the N-linked glycosylation sites may improve accessibility of the protein surface for conjugation to the HIV-1 Env fusion peptide.

In some embodiments, the heterologous carrier protein is a Cholera Toxin B Subunit comprising amino acid substitutions to remove one or more N-linked glycosylation sites. It is believed that removal of the N-linked glycosylation sites may improve accessibility of the protein surface for conjugation to the HIV-1 Env fusion peptide. An exemplary Cholera Toxin B Subunit sequence with modifications to remove one or more N-linked glycosylation sites is provided as:

NGTPQNITDLCAEYHNTQIHTLNDKIFSYTESLAGKREMAIITFKNGATFQVEVPGSQHIDSQKKAIERMKDTLRIAYLT
EAKVEKLCVWNNKTPHAIAAISMAN (SEQ ID NO: 224)

In some embodiments, the heterologous carrier protein is a Cholera Toxin B Subunit comprising amino acid substitutions to remove one or more N-linked glycosylation sites and/or to introduce lysine residues at surface exposed positions of the carrier. It is believed that removal of the N-linked glycosylation sites may improve accessibility of the protein surface for conjugation to the HIV-1 Env fusion peptide.

Any one of the above disclosed heterologous carrier proteins can be fused to any one of the self-assembling protein nanoparticle subunits in the fusion protein of the self-assembling protein nanoparticle carrier.

3. Linker

The heterologous carrier protein fused to the self-assembling protein nanoparticle subunit can be direct linked (for example, the C-terminus of the heterologous carrier protein is linked to the N-terminus of the self-assembling protein nanoparticle subunit by a peptide bond), or indirectly linked by a peptide linker (for example, the C-terminus of the heterologous carrier protein is directly linked to the N-terminus of a peptide linker by a peptide bond, and the C-terminus of the peptide linker is directly linked to the N-terminus of the self-assembling protein nanoparticle subunit by a peptide bond). Any suitable linker can be used to fuse the heterologous carrier protein and the self-assembling protein nanoparticle. In some embodiments, the linker comprises a camel IgG2a hinge (referred to as caIgG2a, EPKIPQPQPKPQPQPQPQPKPQPKPEPE, SEQ ID NO: 327). In some embodiments, the linker comprises a CD8 hinge region, such as KPTTTPAPRPPTPAPTIASQPLSLRPEACRPAAGGAVHTRGLDFACD (SEQ ID NO: 328). In some embodiments, the linker comprises an antibody hinge sequence, such as ggsgEPKSDKTHTPPPAPELLgsgEPKSDKTHTPPPAPELLgsgg (SEQ ID NO: 329). In some embodiments, the linker comprises a flexible protein sequence, such as a glycine serine linker sequence, for example, GGGGSGGGGS (SEQ ID NO: 330).

The linker fusing the carrier protein and the self-assembling nanoparticle subunit can be any suitable length; in some embodiments, the linker is from 10-00 amino acids in length, such as from 10-50 amino acids in length.

In some embodiments, the fusion protein comprises or consists of a rTT carrier protein such as SEQ ID NO: 44 linked to any one of the self-assembling protein nanoparticle subunits provided herein by a peptide linker, such as a caIgG2a linker (e.g., SEQ ID NO: 327), a CD8 linker (e.g., SEQ ID NO: 328), an antibody hinge linker (e.g., SEQ ID NO: 329), or a flexible linker such as a glycine-serine linker (e.g., SEQ ID NO: 330).

In some embodiments, the fusion protein comprises or consists of a rTT carrier protein such as SEQ ID NO: 45 linked to any one of the self-assembling protein nanoparticle subunits provided herein by a peptide linker, such as a caIgG2a linker (e.g., SEQ ID NO: 327), a CD8 linker (e.g., SEQ ID NO: 328), an antibody hinge linker (e.g., SEQ ID NO: 329), or a flexible linker such as a glycine-serine linker (e.g., SEQ ID NO: 330).

In some embodiments, the fusion protein comprises or consists of a rTT carrier protein such as SEQ ID NO: 46 linked to any one of the self-assembling protein nanoparticle subunits provided herein by a peptide linker, such as a caIgG2a linker (e.g., SEQ ID NO: 327), a CD8 linker (e.g., SEQ ID NO: 328), an antibody hinge linker (e.g., SEQ ID NO: 329), or a flexible linker such as a glycine-serine linker (e.g., SEQ ID NO: 330).

In some embodiments, the fusion protein comprises or consists of a rTT carrier protein such as SEQ ID NO: 47 linked to any one of the self-assembling protein nanoparticle subunits provided herein by a peptide linker, such as a caIgG2a linker (e.g., SEQ ID NO: 327), a CD8 linker (e.g., SEQ ID NO: 328), an antibody hinge linker (e.g., SEQ ID NO: 329), or a flexible linker such as a glycine-serine linker (e.g., SEQ ID NO: 330).

In some embodiments, the fusion protein comprises or consists of a rTT carrier protein such as SEQ ID NO: 48 linked to any one of the self-assembling protein nanoparticle subunits provided herein by a peptide linker, such as a caIgG2a linker (e.g., SEQ ID NO: 327), a CD8 linker (e.g., SEQ ID NO: 328), an antibody hinge linker (e.g., SEQ ID NO: 329), or a flexible linker such as a glycine-serine linker (e.g., SEQ ID NO: 330).

In some embodiments, the fusion protein comprises or consists of a rTT carrier protein such as SEQ ID NO: 49 linked to any one of the self-assembling protein nanoparticle subunits provided herein by a peptide linker, such as a caIgG2a linker (e.g., SEQ ID NO: 327), a CD8 linker (e.g., SEQ ID NO: 328), an antibody hinge linker (e.g., SEQ ID NO: 329), or a flexible linker such as a glycine-serine linker (e.g., SEQ ID NO: 330).

In some embodiments, the fusion protein comprises or consists of a rTT carrier protein such as SEQ ID NO: 50 linked to any one of the self-assembling protein nanoparticle subunits provided herein by a peptide linker, such as a caIgG2a linker (e.g., SEQ ID NO: 327), a CD8 linker (e.g., SEQ ID NO: 328), an antibody hinge linker (e.g., SEQ ID NO: 329), or a flexible linker such as a glycine-serine linker (e.g., SEQ ID NO: 330).

In some embodiments, the fusion protein comprises or consists of a rTT carrier protein such as SEQ ID NO: 51 linked to any one of the self-assembling protein nanoparticle subunits provided herein by a peptide linker, such as a caIgG2a linker (e.g., SEQ ID NO: 327), a CD8 linker (e.g., SEQ ID NO: 328), an antibody hinge linker (e.g., SEQ ID NO: 329), or a flexible linker such as a glycine-serine linker (e.g., SEQ ID NO: 330).

In some embodiments, the fusion protein comprises or consists of a rTT carrier protein such as SEQ ID NO: 52 linked to any one of the self-assembling protein nanoparticle subunits provided herein by a peptide linker, such as a caIgG2a linker (e.g., SEQ ID NO: 327), a CD8 linker (e.g., SEQ ID NO: 328), an antibody hinge linker (e.g., SEQ ID NO: 329), or a flexible linker such as a glycine-serine linker (e.g., SEQ ID NO: 330).

In some embodiments, the fusion protein comprises or consists of a rTT carrier protein such as SEQ ID NO: 53 linked to any one of the self-assembling protein nanoparticle subunits provided herein by a

peptide linker, such as a caIgG2a linker (e.g., SEQ ID NO: 327), a CD8 linker (e.g., SEQ ID NO: 328), an antibody hinge linker (e.g., SEQ ID NO: 329), or a flexible linker such as a glycine-serine linker (e.g., SEQ ID NO: 330).

In some embodiments, the fusion protein comprises or consists of a rTT carrier protein such as SEQ ID NO: 54 linked to any one of the self-assembling protein nanoparticle subunits provided herein by a peptide linker, such as a caIgG2a linker (e.g., SEQ ID NO: 327), a CD8 linker (e.g., SEQ ID NO: 328), an antibody hinge linker (e.g., SEQ ID NO: 329), or a flexible linker such as a glycine-serine linker (e.g., SEQ ID NO: 330).

In some embodiments, the fusion protein comprises or consists of a rTT carrier protein such as SEQ ID NO: 55 linked to any one of the self-assembling protein nanoparticle subunits provided herein by a peptide linker, such as a caIgG2a linker (e.g., SEQ ID NO: 327), a CD8 linker (e.g., SEQ ID NO: 328), an antibody hinge linker (e.g., SEQ ID NO: 329), or a flexible linker such as a glycine-serine linker (e.g., SEQ ID NO: 330).

In some embodiments, the fusion protein comprises or consists of a rTT carrier protein such as SEQ ID NO: 56 linked to any one of the self-assembling protein nanoparticle subunits provided herein by a peptide linker, such as a caIgG2a linker (e.g., SEQ ID NO: 327), a CD8 linker (e.g., SEQ ID NO: 328), an antibody hinge linker (e.g., SEQ ID NO: 329), or a flexible linker such as a glycine-serine linker (e.g., SEQ ID NO: 330).

In some embodiments, the fusion protein comprises or consists of a rTT carrier protein such as SEQ ID NO: 57 linked to any one of the self-assembling protein nanoparticle subunits provided herein by a peptide linker, such as a caIgG2a linker (e.g., SEQ ID NO: 327), a CD8 linker (e.g., SEQ ID NO: 328), an antibody hinge linker (e.g., SEQ ID NO: 329), or a flexible linker such as a glycine-serine linker (e.g., SEQ ID NO: 330).

In some embodiments, the fusion protein comprises or consists of a rTT carrier protein such as SEQ ID NO: 58 linked to any one of the self-assembling protein nanoparticle subunits provided herein by a peptide linker, such as a caIgG2a linker (e.g., SEQ ID NO: 327), a CD8 linker (e.g., SEQ ID NO: 328), an antibody hinge linker (e.g., SEQ ID NO: 329), or a flexible linker such as a glycine-serine linker (e.g., SEQ ID NO: 330).

In some embodiments, the fusion protein comprises or consists of a rTT carrier protein such as SEQ ID NO: 59 linked to any one of the self-assembling protein nanoparticle subunits provided herein by a peptide linker, such as a caIgG2a linker (e.g., SEQ ID NO: 327), a CD8 linker (e.g., SEQ ID NO: 328), an antibody hinge linker (e.g., SEQ ID NO: 329), or a flexible linker such as a glycine-serine linker (e.g., SEQ ID NO: 330).

In some embodiments, the fusion protein comprises or consists of a rTT carrier protein such as SEQ ID NO: 60 linked to any one of the self-assembling protein nanoparticle subunits provided herein by a peptide linker, such as a caIgG2a linker (e.g., SEQ ID NO: 327), a CD8 linker (e.g., SEQ ID NO: 328), an

antibody hinge linker (e.g., SEQ ID NO: 329), or a flexible linker such as a glycine-serine linker (e.g., SEQ ID NO: 330).

In some embodiments, the fusion protein comprises or consists of a rTT carrier protein such as SEQ ID NO: 61 linked to any one of the self-assembling protein nanoparticle subunits provided herein by a peptide linker, such as a caIgG2a linker (e.g., SEQ ID NO: 327), a CD8 linker (e.g., SEQ ID NO: 328), an antibody hinge linker (e.g., SEQ ID NO: 329), or a flexible linker such as a glycine-serine linker (e.g., SEQ ID NO: 330).

In some embodiments, the fusion protein comprises or consists of a rTT carrier protein such as SEQ ID NO: 62 linked to any one of the self-assembling protein nanoparticle subunits provided herein by a peptide linker, such as a caIgG2a linker (e.g., SEQ ID NO: 327), a CD8 linker (e.g., SEQ ID NO: 328), an antibody hinge linker (e.g., SEQ ID NO: 329), or a flexible linker such as a glycine-serine linker (e.g., SEQ ID NO: 330).

In some embodiments, the fusion protein comprises or consists of a rTT carrier protein such as SEQ ID NO: 63 linked to any one of the self-assembling protein nanoparticle subunits provided herein by a peptide linker, such as a caIgG2a linker (e.g., SEQ ID NO: 327), a CD8 linker (e.g., SEQ ID NO: 328), an antibody hinge linker (e.g., SEQ ID NO: 329), or a flexible linker such as a glycine-serine linker (e.g., SEQ ID NO: 330).

In some embodiments, the fusion protein comprises or consists of a rTT carrier protein such as SEQ ID NO: 64 linked to any one of the self-assembling protein nanoparticle subunits provided herein by a peptide linker, such as a caIgG2a linker (e.g., SEQ ID NO: 327), a CD8 linker (e.g., SEQ ID NO: 328), an antibody hinge linker (e.g., SEQ ID NO: 329), or a flexible linker such as a glycine-serine linker (e.g., SEQ ID NO: 330).

In some embodiments, the fusion protein comprises or consists of a HiD carrier protein such as SEQ ID NO: 65 linked to any one of the self-assembling protein nanoparticle subunits provided herein by a peptide linker, such as a caIgG2a linker (e.g., SEQ ID NO: 327), a CD8 linker (e.g., SEQ ID NO: 328), an antibody hinge linker (e.g., SEQ ID NO: 329), or a flexible linker such as a glycine-serine linker (e.g., SEQ ID NO: 330).

In some embodiments, the fusion protein comprises or consists of a CRM197 carrier protein such as SEQ ID NO: 66 linked to any one of the self-assembling protein nanoparticle subunits provided herein by a peptide linker, such as a caIgG2a linker (e.g., SEQ ID NO: 327), a CD8 linker (e.g., SEQ ID NO: 328), an antibody hinge linker (e.g., SEQ ID NO: 329), or a flexible linker such as a glycine-serine linker (e.g., SEQ ID NO: 330).

In some embodiments, the fusion protein comprises or consists of a CRM197 carrier protein such as SEQ ID NO: 67 linked to any one of the self-assembling protein nanoparticle subunits provided herein by a peptide linker, such as a caIgG2a linker (e.g., SEQ ID NO: 327), a CD8 linker (e.g., SEQ ID NO: 328), an antibody hinge linker (e.g., SEQ ID NO: 329), or a flexible linker such as a glycine-serine linker (e.g., SEQ ID NO: 330).

In some embodiments, the fusion protein comprises or consists of a OMPC carrier protein such as SEQ ID NO: 222 linked to any one of the self-assembling protein nanoparticle subunits provided herein by a peptide linker, such as a caIgG2a linker (e.g., SEQ ID NO: 327), a CD8 linker (e.g., SEQ ID NO: 328), an antibody hinge linker (e.g., SEQ ID NO: 329), or a flexible linker such as a glycine-serine linker (e.g., SEQ ID NO: 330).

In some embodiments, the fusion protein comprises or consists of a Outer-membrane lipoprotein carrier protein such as SEQ ID NO: 223 linked to any one of the self-assembling protein nanoparticle subunits provided herein by a peptide linker, such as a caIgG2a linker (e.g., SEQ ID NO: 327), a CD8 linker (e.g., SEQ ID NO: 328), an antibody hinge linker (e.g., SEQ ID NO: 329), or a flexible linker such as a glycine-serine linker (e.g., SEQ ID NO: 330).

In some embodiments, the fusion protein comprises or consists of a Cholera Toxin B Subunit carrier protein such as SEQ ID NO: 224 linked to any one of the self-assembling protein nanoparticle subunits provided herein by a peptide linker, such as a caIgG2a linker (e.g., SEQ ID NO: 327), a CD8 linker (e.g., SEQ ID NO: 328), an antibody hinge linker (e.g., SEQ ID NO: 329), or a flexible linker such as a glycine-serine linker (e.g., SEQ ID NO: 330).

4. Heterologous T-cell helper epitope

In some embodiments, the fusion protein further comprises a heterologous T-cell helper epitope sequence. It is believed that the presence of the heterologous T-cell helper epitope sequence on the self-assembling protein nanoparticle carrier will improve the immune response elicited by an immunogenic conjugate containing the carrier conjugated to HIV-1 Env fusion peptides as disclosed herein. Any suitable heterologous T-cell helper epitope sequence can be included on the fusion protein. In some embodiments, the amino acid sequence of the T-cell helper epitope is the sequence of a pan DR epitope (PADRE), such as AKFVAAWTLKAAA (SEQ ID NO: 221). In some embodiments, the amino acid sequence of the T-cell helper epitope is the sequence of a P2 epitope, such as QYIKANSKFIGITEL (SEQ ID NO: 68). In some embodiments, the amino acid sequence of the T-cell helper epitope is the sequence of a TpD epitope, such as ILMQYIKANSKFIGKVSVRQSIALSSLMVAQ (SEQ ID NO: 69). In some embodiments, the amino acid sequence of the T-cell helper epitope is the sequence of an HIV-1 Env epitope, such as HIV-1 Env residues 31-45 according to the HXB2 numbering system, for example, AENLWVTVYYGVPVW (SEQ ID NO: 70) or TEKLWVTVYYGVPVW (SEQ ID NO: 71). In some embodiemnts, the amino acid sequence of the T-cell helper epitope is selected from any one of:

(a) SEQ ID NO: 67;

(b) SEQ ID NO: 68;

(c) SEQ ID NO: 69;

(d) the sequence of HIV-1 Env residues 31-45 according to the HXB2 numbering system (Env31-45 epitope); or

(e) a combination of any one of (a) and (b); (a) and (c); (a) and (d); (b) and (c); (b) and (d); (c) and (d); (a), (b), and (c); (a), (b), and (d); (a), (c), and (d); (b), (c), and (d); or (a), (b), (c), and (d).

In some embodiments, the amino acid sequence of the T-cell helper epitope is the sequence of a p458m epitope, such as NEDQKIGIEIIKRALKI (SEQ ID NO: 225). In some embodiments, the amino acid sequence of the T-cell helper epitope is the sequence of a P30 epitope, such as FNNFTVSFWLRVPKVSASHLE (SEQ ID NO: 226). In some embodiments, the amino acid sequence of the T-cell helper epitope is the sequence of a diphtheria toxin epitope, such as PVFAGANYAAWAVNVAQVI (DTD271–290, SEQ ID NO: 227), HHNTEEIVAQSIALSSLMV (DTD321–340, SEQ ID NO: 228), QSIALSSLMVAQAIPLVGEL (DTD331–350, SEQ ID NO: 229), VDIGFAAYNFVESIINLFQV (DTD351–370, SEQ ID NO: 230), QGESGHDIKITAENTPLPIA (DTD411–430, SEQ ID NO: 231), or GVLLPTIPGKLDVNKSKTHI (DTD431–450, SEQ ID NO: 232). In some embodiments, the amino acid sequence of the T-cell helper epitope is the sequence of a tetanus toxin epitope, such as NSVDDALINSTKIYSYFPSV (TT 580-599, SEQ ID NO: 233), QYIKANSKFIGITEL (TT 830-844, SEQ ID NO: 234), PGINGKAIHLVNNESSE (TT, 916-932, SEQ ID NO: 235), FNNFTVSFWLRVPKVSASHLE (TT, 947-967, SEQ ID NO: 236). In some embodiments, the amino acid sequence of the T-cell helper epitope is the sequence of an HIV-1 Env epitope, such as TQLFNSTWFNSTWST (HIV-1 Env 388-402, SEQ ID NO: 237), EQIWNHTTWMEWDRE (HIV-1 Env 620-634, SEQ ID NO: 238), IRGQIRCSSNITGLLLTRDGGNNAAA (HIV_env_DRBO101_1, SEQ ID NO: 239), QCTHGIRPVVSTQLLLNGSLAEE (HIV_env_DRBO101_2, SEQ ID NO: 240), NDNTSYRLISCNTSVITQACPKV (HIV_env_DRBO101_3, SEQ ID NO: 241), SENFTNNAKIIIVQLNESVVINC (HIV_env_DRBO101_5, SEQ ID NO: 242), EVVIRSENFTNNAKTIIVQLNES (HIV_env_DRBO101_7, SEQ ID NO: 243), TVQCTHGIRPVVSTQLLLNGSLA (HIV_env_DRBO101_11, SEQ ID NO: 244), or ESVVINCTRPNNNTRRSIHIGPG (HIV_env_DRBO101_14, SEQ ID NO: 245).

The heterologous T-cell helper epitope can be located at any suitable section of the fusion protein, including (but not limited to) the N-terminus, the C-terminus, and between the heterologous carrier protein and the self-assembling protein nanoparticle subunit. In some embodiments, the heterologous T-cell helper epitope is separated from the carrier protein and/or the self-assembling protein nanoparticle subunit in the fusion protein by one or more peptide linkers.

5. Targeting Moiety

In some embodiments, the immunogenic conjugate further includes a moiety that targets the immune system in a subject to enhance the immune response to the HIV-1 Env fusion peptide on the immunogenic conjugate. The moiety can be, for example, a moiety that binds to components of the immune system in the subject, such as a pattern recognition receptor, a dendritic cell, or to antigens located in B-cell developmental regions of the immune system, such as germinal centers.

In some embodiments, the fusion protein is linked to a moiety that specifically binds to a pattern recognition receptor agonist, such as a toll-like receptor (TLR) agonist, a Stimulator of Interferon Genes (STING) agonist, a C-type lectin receptor (CLR) agonist, a RIG-I-like receptor (RLR) agonist, or a NOD-like receptor (NLR) agonist.

In several embodiments, the moiety can be a pattern recognition receptor agonist. Non-limiting examples of pattern recognition receptor agonists include TLR-1/2/6 agonists (*e.g.*, lipopeptides and glycolipids, such as Pam2cys or Pam3cys lipopeptides); TLR-3 agonists (e.g., dsRNA, such as PolyI:C, and nucleotide base analogs); TLR-4 agonist (*e.g.*, lipopolysaccharide (LPS) derivatives and small molecule analogs of pyrimidoindole); TLR5 agonists (*e.g.*, Flagellin); TLR-7/8 agonists (*e.g.*, ssRNA and nucleotide base analogs, including derivatives of imidazoquinolines, hydroxy-adenine, benzonapthyridine and loxoribine); and TLR-9 agonists (*e.g.*, unmethylated CpG); Stimulator of Interferon Genes (STING) agonists (*e.g.*, cyclic dinucleotides, such as cyclic diadenylate monophosphate); C-type lectin receptor (CLR) agonists (such as various mono, di, tri and polymeric sugars that can be linear or branched, *e.g.*, mannose, Lewis-X tri-saccharides, etc.); RIG-I-like receptor (RLR) agonists; and NOD-like receptor (NLR) agonists (such as peptidogylcans and structural motifs from bacteria, *e.g.*, meso-diaminopimelic acid and muramyl dipeptide); and combinations thereof. In several embodiments, the pattern recognition receptor agonist can be a TLR agonist, such as an imidazoquinoline-based TLR-7/8 agonist. For example, the adjuvant can be Imiquimod (R837) or Resiquimod (R848), which are approved by the FDA for human use.

In several embodiments, the moiety can be a TLR-7 agonist, a TLR-8 agonist and/or a TLR-7/8 agonist. Numerous such agonists are known, including many different imidazoquinoline compounds. Imidazoquinolines are synthetic immunomodulatory drugs that act by binding Toll-like receptors 7 and 8 (TLR-7/TLR-8) on antigen presenting cells (*e.g.*, dendritic cells), structurally mimicking these receptors' natural ligand, viral single-stranded RNA. Imidazoquinolines are heterocyclic compounds comprising a fused quinoline-imidazole skeleton. Derivatives, salts (including hydrates, solvates, and N-oxides), and prodrugs thereof also are contemplated by the present disclosure. Particular imidazoquinoline compounds are known in the art, see for example, U.S. Patent No. 6,518,265; and U.S. Patent No. 4,689,338. In some non-limiting embodiments, the imidazoquinoline compound is not imiquimod and/or is not resiquimod.

The moiety that targets the immune system in a subject can be linked to the immunogenic conjugate by any suitable means.

In some embodiments, the fusion protein includes the sequence of flagellin subunit.

In some embodiments, the fusion protein of the self-assembling protein-nanoparticle carrier includes a streptavidin sequence, and the moiety that targets the immune system in a subject is biotinylated, for example, a biotinylated pattern recognition receptor agonist, such as a biotinylated TLR agonist, a biotinylated STING agonist, a biotinylated CLR agonist, a biotinylated RLR agonist, or a biotinylated NLR agonist. The biotinylated moiety can be linked to the self-assembling protein-nanoparticle carrier.

In some embodiments, the moiety that targets the immune system in the subject is conjugated to the self-assembling protein nanoparticle carrier using the same conjugate method as that used to conjugate the

HIV-1 Env fusion peptide to the self-assembling protein nanoparticle carrier. In some such embodiments, conjugation of the moiety that targets the immune system and the HIV-1 fusion peptide to the self-assembling protein nanoparticle carrier can be completed in the same reaction. For example, both the moiety that targets the immune system in the subject and the HIV-1 Env fusion peptide can be linked to a cysteine residue for conjugation to the self-assembling protein nanoparticle carrier as described herein. In some embodiments, the HIV-1 Env fusion peptide linked to a cysteine residue is mixed with a small amount of TLR-7 or 8 agonist modified to include a reactive –SH group, so that both the HIV-1 Env fusion peptide and the TLR7/8-agonist are conjugated via a single reaction to a bifunctional crosslinker-activated self-assembled protein nanoparticle carrier.

6. Exemplary fusion protein embodiments

In several embodiments, the fusion proteins of the self-assembling protein nanoparticle carrier comprise, consist essentially of, or consist of the amino acid sequence of any one of fusion proteins listed in the following table (showing SEQ ID NOs: 72-219, 246-257, and 331-397), or an amino acid sequence at least 90% (such as at least 95%) identical to any one of SEQ ID NOs: 72-219, 246-257, or 331-397 that self-assembles into a protein nanoparticle under suitable conditions. In some embodiments, the fusion proteins of the self-assembling protein nanoparticle carrier comprise, consist essentially of, or consist of the amino acid sequence set forth as any one of SEQ ID NOs: 73, 76, 79, 100, 101, 109, 116, 167, 172, 180, 197, or 211, or an amino acid sequence at least 90% (such as at least 95%) identical to any one of SEQ ID NOs: 73, 76, 79, 100, 101, 109, 116, 167, 172, 180, 197, or 211 that self-assembles into a protein nanoparticle under suitable conditions.

Table 1. Exemplary sequences of fusion proteins containing a protein nanoparticle subunit fused to a heterologous carrier protein and optionally a heterologous T-cell helper epitope.

SEQ ID NO	Name	Sequence
		Lumazine Synthase
72	LS-20-CRM	*QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAA GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTLEQAI ERAGTKHGNKGWEAALSAIEMANLFKSLR*ggksggnkksdgvkessesgGADDVVDSSKSFV MENFSSYHGTKPGYVDSIQKGIQKPKSGTQGNYDDDWKEFYSTDNKYDAAGYSVDNENPLSG KAGGVVKVTYPGLTKVLALKVDNAETIKKELGLSLTEPLMEQVGTEEFIKRFGDGASRVVLS LPFAEGSSSVEYINNWEQAKALSVELEINFETRGKRGQDAMYEYMAQACAGNRVRRSVGSSL SCINLDWDVIRDKTKTKIESLKEHGPIKNKMSESPNKTVSEEKAKQYLEEFHQTALEHPELS ELKTVTGTNPVFAGANYAAWAVNVAQVIDSETADNLEKTTAALSILPGIGSVMGIADGAVHH NTEEIVAQSIALSSLMVAQAIPLVGELVDIGFAAYNFVESIINLFQVVHNSYNRPAYSPGHK TQPFLHDGYAVSWNTVEDSIIRTGFQGESGHDIKITAENTPLPIAGVLLPTIPGKLDVNKSK THISVNGRKIRMRCRAIDGDVTFCRPKSPVYVGNGVHANLHVAFHRSSSEKIHSNEISSDSI GVLGYQKTVDHTKVNSKLSLFFEIKS
73	LS-20-rTT	*QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAA GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTLEQAI ERAGTKHGNKGWEAALSAIEMANLFKSLR*ggksggnkksdgvkessesgMKNLDCWVDNEED IDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAIHLVNNESSEVIVHKA MDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNL IWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGVLMGSAEITGLG

		AIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITFLRDFWGNPLRY DTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNEI DSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDL KTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFVPTDE GWTND
74	LS-20-HID	*QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAA* *GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTLEQAI* *ERAGTKHGNKGWEAALSAIEMANLFKSLR*ggksggnkksdgvkessesgSNMANTQMKSDKI IIAHRGASGYLPEHTLESKALAFAQQADYLEQDLAMTKDGRLVVIHDHFLDGLTDVAKKFPH RHRKDGRYYVIDFTLKEIQSLEMTENFETKDGKQAQVYPNRFPLWKSHFRIHTFEDEIEFIQ GLEKSTGKKVGIYPEIKAPWFHHQNGKDIAAETLKVLKKYGYDKKTDMVYLQTFDFNELKRI KTELLPQMGMDLKLVQLIAYTDWKETQEKDPKGYWVNYNYDWMFKPGAMAEVVKYADGVGPG WYMLVNKEESKPDNIVYTPLVKELAQYNVEVHPYTVRKDALPEFFTDVNQMYDALLNKSGAT GVFTDFPDTGVEFLKGIK
75	LS-PADRE-Env31-CRM	*QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAA* *GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTLEQAI* *ERAGTKHGNKGWEAALSAIEMANLFKSLR*SLVR**AKFVAAWTLKAAA**GSLVR**AENLWVTVYYG** **VPVW**slvrgGADDVVDSSKSFVMENFSSYHGTKPGYVDSIQKGIQKPKSGTQGNYDDDWKEF YSTDNKYDAAGYSVDNENPLSGKAGGVVKVTYPGLTKVLALKVDNAETIKKELGLSLTEPLM EQVGTEEFIKRFGDGASRVVLSLPFAEGSSSVEYINNWEQAKALSVELEINFETRGKRGQDA MYEYMAQACAGNRVRRSVGSSLSCINLDWDVIRDKTKTKIESLKEHGPIKNKMSESPNKTVS EEKAKQYLEEFHQTALEHPELSELKTVTGTNPVFAGANYAAWAVNVAQVIDSETADNLEKTT AALSILPGIGSVMGIADGAVHHNTEEIVAQSIALSSLMVAQAIPLVGELVDIGFAAYNFVES IINLFQVVHNSYNRPAYSPGHKTQPFLHDGYAVSWNTVEDSIIRTGFQGESGHDIKITAENT PLPIAGVLLPTIPGKLDVNKSKTHISVNGRKIRMRCRAIDGDVTFCRPKSPVYVGNGVHANL HVAFHRSSSEKIHSNEISSDSIGVLGYQKTVDHTKVNSKLSLFFEIKS
76	LS-PADRE-Env31-rTT	*QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAA* *GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTLEQAI* *ERAGTKHGNKGWEAALSAIEMANLFKSLR*SLVRSLVR**AKFVAAWTLKAAA**GSLVR**AENLWVT** **VYYGVPVW**slvrgMKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDA QLVPGINGKAIHLVNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEY SIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITI TNDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNP KEIEKLYTSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNG KLNIYYRRLYNGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLD RILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDIL IASNWYFNHLKDKILGCDWYFVPTDEGWTND
77	LS-PADRE-Env31-HID	*QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAA* *GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTLEQAI* *ERAGTKHGNKGWEAALSAIEMANLFKSLR*SLVRSLVRAKFVAAWTLKAAAGSLVRAENLWVT VYYGVPVWslvrgSNMANTQMKSDKIIIAHRGASGYLPEHTLESKALAFAQQADYLEQDLAM TKDGRLVVIHDHFLDGLTDVAKKFPHRHRKDGRYYVIDFTLKEIQSLEMTENFETKDGKQAQ VYPNRFPLWKSHFRIHTFEDEIEFIQGLEKSTGKKVGIYPEIKAPWFHHQNGKDIAAETLKV LKKYGYDKKTDMVYLQTFDFNELKRIKTELLPQMGMDLKLVQLIAYTDWKETQEKDPKGYWV NYNYDWMFKPGAMAEVVKYADGVGPGWYMLVNKEESKPDNIVYTPLVKELAQYNVEVHPYTV RKDALPEFFTDVNQMYDALLNKSGATGVFTDFPDTGVEFLKGIK
78	rTT_degly-LS	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDgggsgggsgggs*MQIYEGKLTAEGLRFGIVASRFNHALVDRLVE* *GAIDCIVRHGGREEDITLVRVPGSWEIPVAAGELARKEDIDAVIAIGVLIRGATPHFDYIAS* *EVSKGLADLSLELRKPITFGVITADTLEQAIERAGTKHGNKGWEAALSAIEMANLFKSLR*
79	LS-rTT_degly	*QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDCIVRHGGREEDITLVRVPGSWEIPVAA* *GELARKEDIDAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTLEQAI* *ERAGTKHGNKGWEAALSAIEMANLFKSLR*GGSGGGSGGGQMKNLDCWVDNEEDIDVILKKST ILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHLVNNEASEVIVHKAMDIEYNDMF NNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAG EVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGVLMGSAEITGLGAIREDNQIT

		LKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITFLRDFWGNPLRYDTEYYLIPV ASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNEIDSFVKSGDF IKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKL YDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND
80	rTT_degly -10f-LS	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDggggsaeaaakeaaakaggggsggggsggggsggggsggggsgg ggsgggg*MQIYEGKLTAEGLRFGIVASRFNHALVDRLVEGCIDCIVRHGGREEDITLVRVPG SWEIPVAAGELARKEDIDAVIAIGVLIRGATPHFDYIASEVSKGLANLSLELRKPITFGVIT ADTLEQAIERAGTKHGNKCWEAALSAIEMANLFKSLR*
81	rTT_degly -r8-LS	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDaeaaaakeaaakeaaakeaaakaleaeaaaakeaaakeaaakeaaa ka*MQIYEGKLTAEGLRFGIVASRFNHALVDRLVEGCIDCIVRHGGREEDITLVRVPGSWEIP VAAGELARKEDIDAVIAIGVLIRGATPHFDYIASEVSKGLANLSLELRKPITFGVITADTLE QAIERAGTKHGNKCWEAALSAIEMANLFKSLR*
82	rTT_degly -12ap-LS	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDgsapapapapapapapapapapapapapasg*MQIYEGKLTAEGLRF GIVASRFNHALVDRLVEGCIDCIVRHGGREEDITLVRVPGSWEIPVAAGELARKEDIDAVIA IGVLIRGATPHFDYIASEVSKGLANLSLELRKPITFGVITADTLEQAIERAGTKHGNKCWEA ALSAIEMANLFKSLR*
83	rTT_degly -10pa-LS	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDapapapapapapapapapapapap*MQIYEGKLTAEGLRFGIVASRFNH ALVDRLVEGCIDCIVRHGGREEDITLVRVPGSWEIPVAAGELARKEDIDAVIAIGVLIRGAT PHFDYIASEVSKGLANLSLELRKPITFGVITADTLEQAIERAGTKHGNKCWEAALSAIEMAN LFKSLR*
84	rTT_degly -2rf-LS	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDaeaaakeaaakagsgsgs*MQIYEGKLTAEGLRFGIVASRFNHAL VDRLVEGCIDCIVRHGGREEDITLVRVPGSWEIPVAAGELARKEDIDAVIAIGVLIRGATPH FDYIASEVSKGLANLSLELRKPITFGVITADTLEQAIERAGTKHGNKCWEAALSAIEMANLF KSLR*
85	rTT_degly -3f-LS	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN

		GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDggggsggggsggggs*MQIYEGKLTAEGLRFGIVASRFNHALVDR* *LVEGCIDCIVRHGGREEDITLVRVPGSWEIPVAAGELARKEDIDAVIAIGVLIRGATPHFDY* *IASEVSKGLANLSLELRKPITFGVITADTLEQAIERAGTKHGNKCWEAALSAIEMANLFKSL* *R*
86	rTT_degly -5gs-LS	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDgsgsgsgsgsgsasg*MQIYEGKLTAEGLRFGIVASRFNHALVDRLV* *EGCIDCIVRHGGREEDITLVRVPGSWEIPVAAGELARKEDIDAVIAIGVLIRGATPHFDYIA* *SEVSKGLANLSLELRKPITFGVITADTLEQAIERAGTKHGNKCWEAALSAIEMANLFKSLR*
87	rTT_degly -3pa-LS	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDapapapggggsM*QIYEGKLTAEGLRFGIVASRFNHALVDRLVEG* *CIDCIVRHGGREEDITLVRVPGSWEIPVAAGELARKEDIDAVIAIGVLIRGATPHFDYIASE* *VSKGLANLSLELRKPITFGVITADTLEQAIERAGTKHGNKCWEAALSAIEMANLFKSLR*
88	LS-r8- rTT_degly	*MQIYEGKLTAEGLRFGIVASRFNHALVDRLVEGCIDCIVRHGGREEDITLVRVPGSWEIPVA* *AGELARKEDIDAVIAIGVLIRGATPHFDYIASEVSKGLANLSLELRKPITFGVITADTLEQA* *IERAGTKHGNKCWEAALSAIEMANLFKSLR*aeaaakeaaakeaaakeaaakaleaeaaakea aakeaaakeaaakaMKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPD AQLVPGINGKAIHLVNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNE YSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFIT ITNDRLSSANLYINGVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALN PKEIEKLYTSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTN GKLNIYYRRLYNGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNL DRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDI LIASNWYFNHLKDKILGCDWYFVPTDEGWTND
89	LS-8f- rTT_degly	*MQIYEGKLTAEGLRFGIVASRFNHALVDRLVEGCIDCIVRHGGREEDITLVRVPGSWEIPVA* *AGELARKEDIDAVIAIGVLIRGATPHFDYIASEVSKGLANLSLELRKPITFGVITADTLEQA* *IERAGTKHGNKCWEAALSAIEMANLFKSLR*ggggsggggsggggsggggsggggsaeaaake aaakaggggsMKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLV PGINGKAIHLVNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSII SSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITND RLSSANLYINGVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEI EKLYTSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLN IYYRRLYNGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRIL RVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIAS NWYFNHLKDKILGCDWYFVPTDEGWTND
90	LS-r6- rTT_degly	*MQIYEGKLTAEGLRFGIVASRFNHALVDRLVEGCIDCIVRHGGREEDITLVRVPGSWEIPVA* *AGELARKEDIDAVIAIGVLIRGATPHFDYIASEVSKGLANLSLELRKPITFGVITADTLEQA* *IERAGTKHGNKCWEAALSAIEMANLFKSLR*aeaaakeaaakeaaakaleaeaaakeaaakea aakaMKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGK AIHLVNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKH SLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSAN LYINGVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTS YLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRL YNGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNA PGIPLYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNH LKDKILGCDWYFVPTDEGWTND

91	LS-12ap-rTT_degly	*MQIYEGKLTAEGLRFGIVASRFNHALVDRLVEGCIDCIVRHGGREEDITLVRVPGSWEIPVA AGELARKEDIDAVIAIGVLIRGATPHFDYIASEVSKGLANLSLELRKPITFGVITADTLEQA IERAGTKHGNKCWEAALSAIEMANLFKSLR*gsapapapapapapapapapapapapasgMKN LDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHLVNN EASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSG WSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGVL MGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITFL RDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFI IKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYK KMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILG CDWYFVPTDEGWTND
92	LS-10pa-rTT_degly	*MQIYEGKLTAEGLRFGIVASRFNHALVDRLVEGCIDCIVRHGGREEDITLVRVPGSWEIPVA AGELARKEDIDAVIAIGVLIRGATPHFDYIASEVSKGLANLSLELRKPITFGVITADTLEQA IERAGTKHGNKCWEAALSAIEMANLFKSLR*apapapapapapapapapMKNLDCWVDNEE DIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHLVNNEASEVIVHK AMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNN LIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGVLMGSAEITGL GAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITFLRDFWGNPLR YDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNE IDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRD LKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFVPTD EGWTND
93	LS-f2r-rTT_degly	*MQIYEGKLTAEGLRFGIVASRFNHALVDRLVEGCIDCIVRHGGREEDITLVRVPGSWEIPVA AGELARKEDIDAVIAIGVLIRGATPHFDYIASEVSKGLANLSLELRKPITFGVITADTLEQA IERAGTKHGNKCWEAALSAIEMANLFKSLR*gsgsgaeaaakeaaakaMKNLDCWVDNEEDID VILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHLVNNEASEVIVHKAMD IEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIW TLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGVLMGSAEITGLGAI REDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITFLRDFWGNPLRYDT EYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNEIDS FVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDLKT YSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGW TND
94	LS-3f-rTT_degly	*MQIYEGKLTAEGLRFGIVASRFNHALVDRLVEGCIDCIVRHGGREEDITLVRVPGSWEIPVA AGELARKEDIDAVIAIGVLIRGATPHFDYIASEVSKGLANLSLELRKPITFGVITADTLEQA IERAGTKHGNKCWEAALSAIEMANLFKSLR*ggggsggggsggggsMKNLDCWVDNEEDIDVI LKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHLVNNEASEVIVHKAMDIE YNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIWTL KDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGVLMGSAEITGLGAIRE DNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITFLRDFWGNPLRYDTEY YLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNEIDSFV KSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDLKTYS VQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTN D
95	LS-5gs-rTT_degly	*MQIYEGKLTAEGLRFGIVASRFNHALVDRLVEGCIDCIVRHGGREEDITLVRVPGSWEIPVA AGELARKEDIDAVIAIGVLIRGATPHFDYIASEVSKGLANLSLELRKPITFGVITADTLEQA IERAGTKHGNKCWEAALSAIEMANLFKSLR*gsgsgsgsgsasgMKNLDCWVDNEEDIDVILK KSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHLVNNEASEVIVHKAMDIEYN DMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKD SAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGVLMGSAEITGLGAIREDN QITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITFLRDFWGNPLRYDTEYYL IPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNEIDSFVKS GDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQ LKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND
96	LS-3pa-rTT_degly	*MQIYEGKLTAEGLRFGIVASRFNHALVDRLVEGCIDCIVRHGGREEDITLVRVPGSWEIPVA AGELARKEDIDAVIAIGVLIRGATPHFDYIASEVSKGLANLSLELRKPITFGVITADTLEQA IERAGTKHGNKCWEAALSAIEMANLFKSLR*ggggspapapasgMKNLDCWVDNEEDIDVILK KSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHLVNNEASEVIVHKAMDIEYN DMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKD SAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGVLMGSAEITGLGAIREDN QITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITFLRDFWGNPLRYDTEYYL IPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNEIDSFVKS

		GDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQ LKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND
97	rTT-P2-LS-Padre	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAIHL VNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDsgsa*GKKGSSQYIKANSKFIGITEL**MQIYEGKLTAEGLRFGIVA* *SRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAAGELARKENISAVIAIGVL* *IRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTLEQAIERAGTKHGNKGWEAALSA* *IEMANLFKSLR*AKFVAAWTLKAAA
98	rTT-linker-LS-Padre	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAIHL VNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDsgsa*MQIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVR* *HGGREEDITLVRVPGSWEIPVAAGELARKENISAVIAIGVLIRGATPHFDYIASEVSKGLAD* *LSLELRKPITFGVITADTLEQAIERAGTKHGNKGWEAALSAIEMANLFKSLR*AKFVAAWTLK AAA
99	rTT-alphaLinker-LS	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAIHL VNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDsgsa**KALEAQKQK***MQIYEGKLTAEGLRFGIVASRFNHALVDRLV* *EGAIDAIVRHGGREEDITLVRVPGSWEIPVAAGELARKENISAVIAIGVLIRGATPHFDYIA* *SEVSKGLADLSLELRKPITFGVITADTLEQAIERAGTKHGNKGWEAALSAIEMANLFKSLR*
100	LS-alphaLinker-rTT	*MQIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVA* *AGELARKENISAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTLEQA* *IERAGTKHGNKGWEAALSAIEMANLFKSLR*sgsa**KALEAQKQK**MKNLDCWVDNEEDIDVILK KSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAIHLVNNESSEVIVHKAMDIEYN DMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKD SAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGVLMGSAEITGLGAIREDN NITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITFLRDFWGNPLRYDTEYYL IPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNEIDSFVKS GDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQ LKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND
101	rTT-LS-PADRE-SaTyflagellin-CC-YW	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAIHL VNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TTTMLATRNFSGGKSGGNKKSDGVKESSESTNTTIEDED*mqiyegkltaeglrfgivasrfn* *halvdrlvegaidaivrhggreeditlvrvpgsweipvaagelarke**N**i**S**aviaigvlir*gL EVLFQGPGAKFVAAWTLKAAAGDEVDatphfdyiasevskgladlslelrkpitfgvitadt leqaieragtkhgnkgweaalsaiemanlfkslrGGKSGGNKKSDGVNLTDLGLTQSNIQKL DIDITEGDNAGVQITLTNDQALVKVGNFQTQGSVRDIENLRQTIEAQISDLDSQSNTSNASQ VALERVRQLNNNIENLAGETTQAISIGDNANRSAQTLGKINATFRNAIAQGAADDKASNIRL GSSLREIATGLASQSKNLNNQTLLSLSNTNIVQAGSGSARLLSLVNQPVQNAQALVSTGAQQ LIQARSMNSVETAYDSDEIRSRASTLNNVTNGLNTIASNFRNQVAGLDSRLTDVQALAADIK QLPNE
102	rTT-LS-PADRE-SaTyflagellin-CC-YW-degly	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAIHL VNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TTTMLATRNFSGGKSGGNKKSDGVKESSESTNTTIEDED*mqiyegkltaeglrfgivasrfn*

		*halvdrlvegaidaivrhggreeditlvrvpgsweipvaagelarke**N**i**S**aviaigvlir*gL EVLFQGPGAKFVAAWTLKAAAGDEVDatphfdyiasevskgladlslelrkpitfgvitadt leqaieragtkhgnkgweaalsaiemanlfkslrGGKSGGNKKSDGVNLaDLGLTQSNIQKL DIDITEGDNAGVQITLTNDQALVKVGNFQTQGSVRDIENLRQTIEAQISDLDSQSNTSNAaQ VALERVRQLNNNIENLAGETTQAISIGDNANRaAQTLGKINAaFRNAIAQGAADDKASNIRL GSSLREIATGLASQSKNLNNQaLLSLSNTNIVQAGSGSARLLSLVNQPVQNAQALVSTGAQQ LIQARSMNSVETAYDSDEIRSRASTLNNVaNGLNTIASNFRNQVAGLDSRLTDVQALAADIK QLPNE
103	revTT-LS-PADRE-SaTyflagellin-CC-YW	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAIHL VNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TTTMLATRNFSGGKSGGNKKSDGVKESSESTNTTIEDED*mqiyegkltaeglrfgivasrfn* *halvdrlvegaidaivrhggreeditlvrvpgsweipvaagelarke**N**i**S**aviaigvlir*gL EVLFQGPGAKFVAAWTLKAAAGDEVDatphfdyiasevskgladlslelrkpitfgvitadt leqaieragtkhgnkgweaalsaiemanlfkslrGGKSGGNKKSDGVENPLQKIDAALAQVD TLRSDLGAVQNRFNSAITNLGNTVNNLTSARSRIEDSDYATEVSNMSRAQILQQAGTSVLAQ ANQVPQNVLSLLRGSGSAAQVINTNSLSLLTQNNLNKSQSALGTAIERLSSGLRINSAKDDA AGQAIANRFTANIKGLTQASRNANDGISIAQTTEGALNEINNNLQRVRELAVQSANSTNSQS DLDSIQAEITQRLNEIDRVSGQTQFNGVKVLAQDNTLTIQVGANDGETIDIDLKQINSQTLG LDTLN
104	revTT-LS-PADRE-SaTyflagellin-CC-YW-degly	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAIHL VNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TTTMLATRNFSGGKSGGNKKSDGVKESSESTNTTIEDEDgsg*mqiyegkltaeglrfgivas* *rfnhalvdrlvegaidaivrhggreeditlvrvpgsweipvaagelarke**N**i**S**aviaigvli* *rg*LEVLFQGPGAKFVAAWTLKAAAGDEVDatphfdyiasevskgladlslelrkpitfgvit adtleqaieragtkhgnkgweaalsaiemanlfkslrGGKSGGNKKSDGVENPLQKIDAALA QVDTLRSDLGAVQNRFNSAITNLGNTVNNLaSARSRIEDSDYATEVSNMaRAQILQQAGTSV LAQANQVPQNVLSLLRGSGSAAQVINTNSLSLLTQNNLNKaQSALGTAIERLSSGLRINSAK DDAAGQAIANRFTANIKGLTQASRNANDGISIAQTTEGALNEINNNLQRVRELAVQSANSaN SQSDLDSIQAEITQRLNEIDRVSGQTQFNGVKVLAQDNTLTIQVGANDGETIDIDLKQINSQ TLGLDTLN
105	SaTyflagellin-LS-PADRE-rTT-His-CC-YW	NLTDLGLTQSNIQKLDIDITEGDNAGVQITLTNDQALVKVGNFQTQGSVRDIENLRQTIEAQ ISDLDSQSNTSNASQVALERVRQLNNNIENLAGETTQAISIGDNANRSAQTLGKINATFRNA IAQGAADDKASNIRLGSSLREIATGLASQSKNLNNQTLLSLSNTNIVQAGSGSARLLSLVNQ PVQNAQALVSTGAQQLIQARSMNSVETAYDSDEIRSRASTLNNVTNGLNTIASNFRNQVAGL DSRLTDVQALAADIKQLPNEGGKSGGNKKSDGV*mqiyegkltaeglrfgivasrfnhalvdr* *lvegaidaivrhggreeditlvrvpgsweipvaagelarkeNiSaviaigvlir*gLEVLFQG PGAKFVAAWTLKAAAGDEVDatphfdyiasevskgladlslelrkpitfgvitadtleqaie ragtkhgnkgweaalsaiemanlfkslrTTMLATRNFSGGKSGGNKKSDGVKESSESTNTTI EDEDMKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGK AIHLVNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKH SLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSAN LYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTS YLSIT
106	SaTyflagellin-LS-PADRE-rTT-His-CC-YW-degly	NLaDLGLTQSNIQKLDIDITEGDNAGVQITLTNDQALVKVGNFQTQGSVRDIENLRQTIEAQ ISDLDSQSNTSNAaQVALERVRQLNNNIENLAGETTQAISIGDNANRaAQTLGKINAaFRNA IAQGAADDKASNIRLGSSLREIATGLASQSKNLNNQaLLSLSNTNIVQAGSGSARLLSLVNQ PVQNAQALVSTGAQQLIQARSMNSVETAYDSDEIRSRASTLNNVaNGLNTIASNFRNQVAGL DSRLTDVQALAADIKQLPNEGGKSGGNKKSDGVgsg*mqiyegkltaeglrfgivasrfnhal* *vdrlvegaidaivrhggreeditlvrvpgsweipvaagelarkeNiSaviaigvlir*gLEVL FQGPGAKFVAAWTLKAAAGDEVDatphfdyiasevskgladlslelrkpitfgvitadtleq aieragtkhgnkgweaalsaiemanlfkslrTTMLATRNFSGGKSGGNKKSDGVKESSESTN TTIEDEDMKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGI NGKAIHLVNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSM KKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLS SANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKL YTSYLSIT
107	revSaTyflagellin-	ENPLQKIDAALAQVDTLRSDLGAVQNRFNSAITNLGNTVNNLTSARSRIEDSDYATEVSNMS RAQILQQAGTSVLAQANQVPQNVLSLLRGSGSAAQVINTNSLSLLTQNNLNKSQSALGTAIE

	LS-PADRE-rTT-His-CC-YW	RLSSGLRINSAKDDAAGQAIANRFTANIKGLTQASRNANDGISIAQTTEGALNEINNNLQRV RELAVQSANSTNSQSDLDSIQAEITQRLNEIDRVSGQTQFNGVKVLAQDNTLTIQVGANDGE TIDIDLKQINSQTLGLDTLNGGKSGGNKKSDGV*mqiyegkltaeglrfgivasrfnhalvdr* *lvegaidaivrhggreeditlvrvpgsweipvaagelarkeNiSaviaigvlir*gLEVLFQG PGAKFVAAWTLKAAAGDEVDatphfdyiasevskgladlslelrkpitfgvitadtleqaie ragtkhgnkgweaalsaiemanlfkslrTTMLATRNFSGGKSGGNKKSDGVKESSESTNTTI EDEDMKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGK AIHLVNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKH SLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSAN LYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTS YLSIT
108	revSaTyfl agellin-LS-PADRE-rTT-His-CC-YW-degly	ENPLQKIDAALAQVDTLRSDLGAVQNRFNSAITNLGNTVNNLaSARSRIEDSDYATEVSNMa RAQILQQAGTSVLAQANQVPQNVLSLLRGSGSAAQVINTNSLSLLTQNNLNKaQSALGTAIE RLSSGLRINSAKDDAAGQAIANRFTANIKGLTQASRNANDGISIAQTTEGALNEINNNLQRV RELAVQSANSaNSQSDLDSIQAEITQRLNEIDRVSGQTQFNGVKVLAQDNTLTIQVGANDGE TIDIDLKQINSQTLGLDTLNGGKSGGNKKSDGVgsg*mqiyegkltaeglrfgivasrfnhal* *vdrlvegaidaivrhggreeditlvrvpgsweipvaagelarkeNiSaviaigvlir*gLEVL FQGPGAKFVAAWTLKAAAGDEVDatphfdyiasevskgladlslelrkpitfgvitadtleq aieragtkhgnkgweaalsaiemanlfkslrggTTMLATRNFSGGKSGGNKKSDGVKESSES TNTTIEDEDMKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVP GINGKAIHLVNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIIS SMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDR LSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIE KLYTSYLSIT
246	LS-rTT-degly	MQIYEGKLTAEGLRFGIVASRFNHALVDRLVEGCIDCIVRHGGREEDITLVRVPGSWEIPVA AGELARKEDIDAVIAIGVLIRGATPHFDYIASEVSKGLANLSLELRKPITFGVITADTLEQA IERAGTKHGNKCWEAALSAIEMANLFKSLRgsgsgsMKNLDCWVDNEEDIDVILKKSTILNL DINNDIISDISGFNSAVITYPDAQLVPGINGKAIHLVNNEASEVIVHKAMDIEYNDMFNNFT VSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQ ITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGVLMGSAEITGLGAIREDNQITLKLD RCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITFLRDFWGNPLRYDTEYYLIPVASSS KDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNEIDSFVKSGDFIKLY VSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDK QASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND
247	LS-CRM197-degly	MQIYEGKLTAEGLRFGIVASRFNHALVDRLVEGCIDCIVRHGGREEDITLVRVPGSWEIPVA AGELARKEDIDAVIAIGVLIRGATPHFDYIASEVSKGLANLSLELRKPITFGVITADTLEQA IERAGTKHGNKCWEAALSAIEMANLFKSLRgsgsgsDYKDDDDKgsgGADDVVDSSKSFVME NFASYHGTKPGYVDSIQKGIQKPKSGTQGNYDDDWKEFYSTDNKYDAAGYSVDNENPLSGKA GGVVKVTYPGLTKVLALKVDNAETIKKELGLSLTEPLMEQVGTEEFIKRFGDGASRVVLSLP FAEGSSSVEYINNWEQAKALSVELEINFETRGKRGQDAMYEYMAQACAGNRVRRSVGSSLSC INLDWDVIRDKTKTKIESLKEHGPIKNKMSESPNKAVSEEKAKQYLEEFHQTALEHPELSEL KTVTGTNPVFAGANYAAWAVNVAQVIDSETADNLEKTTAALSILPGIGSVMGIADGAVHHNT EEIVAQSIALSSLMVAQAIPLVGELVDIGFAAYNFVESIINLFQVVHNSYNRPAYSPGHKTQ PFLHDGYAVSWNTVEDSIIRTGFQGESGHDIKITAENTPLPIAGVLLPTIPGKLDVNKAKTH ISVNGRKIRMRCRAIDGDVTFCRPKSPVYVGNGVHANLHVAFHRSSSEKIHSNEISSDSIGV LGYQKTVDHTKVNSKLSLFFEIKS
248	LS-2xCRM197-degly	MQIYEGKLTAEGLRFGIVASRFNHALVDRLVEGCIDCIVRHGGREEDITLVRVPGSWEIPVA AGELARKEDIDAVIAIGVLIRGATPHFDYIASEVSKGLANLSLELRKPITFGVITADTLEQA IERAGTKHGNKCWEAALSAIEMANLFKSLRgsgDYKDDDDKgsgGADDVVDSSKSFVMENFA SYHGTKPGYVDSIQKGIQKPKSGTQGNYDDDWKEFYSTDNKYDAAGYSVDNENPLSGKAGGV VKVTYPGLTKVLALKVDNAETIKKELGLSLTEPLMEQVGTEEFIKRFGDGASRVVLSLPFAE GSSSVEYINNWEQAKALSVELEINFETRGKRGQDAMYEYMAQACAGNRVRRSVGSSLSCINL DWDVIRDKTKTKIESLKEHGPIKNKMSESPNKAVSEEKAKQYLEEFHQTALEHPELSELKTV TGTNPVFAGANYAAWAVNVAQVIDSETADNLEKTTAALSILPGIGSVMGIADGAVHHNTEEI VAQSIALSSLMVAQAIPLVGELVDIGFAAYNFVESIINLFQVVHNSYNRPAYSPGHKTQPFL HDGYAVSWNTVEDSIIRTGFQGESGHDIKITAENTPLPIAGVLLPTIPGKLDVNKAKTHISV NGRKIRMRCRAIDGDVTFCRPKSPVYVGNGVHANLHVAFHRSSSEKIHSNEISSDSIGVLGY QKTVDHTKVNSKLSLFFEIKSgggsgggsGADDVVDSSKSFVMENFASYHGTKPGYVDSIQK GIQKPKSGTQGNYDDDWKEFYSTDNKYDAAGYSVDNENPLSGKAGGVVKVTYPGLTKVLALK VDNAETIKKELGLSLTEPLMEQVGTEEFIKRFGDGASRVVLSLPFAEGSSSVEYINNWEQAK ALSVELEINFETRGKRGQDAMYEYMAQACAGNRVRRSVGSSLSCINLDWDVIRDKTKTKIES

		LKEHGPIKNKMSESPNKAVSEEKAKQYLEEFHQTALEHPELSELKTVTGTNPVFAGANYAAW AVNVAQVIDSETADNLEKTTAALSILPGIGSVMGIADGAVHHNTEEIVAQSIALSSLMVAQA IPLVGELVDIGFAAYNFVESIINLFQVVHNSYNRPAYSPGHKTQPFLHDGYAVSWNTVEDSI IRTGFQGESGHDIKITAENTPLPIAGVLLPTIPGKLDVNKAKTHISVNGRKIRMRCRAIDGD VTFCRPKSPVYVGNGVHANLHVAFHRSSSEKIHSNEISSDSIGVLGYQKTVDHTKVNSKLSL FFEIKS
249	LS-HID	MQIYEGKLTAEGLRFGIVASRFNHALVDRLVEGCIDCIVRHGGREEDITLVRVPGSWEIPVA AGELARKEDIDAVIAIGVLIRGATPHFDYIASEVSKGLANLSLELRKPITFGVITADTLEQA IERAGTKHGNKCWEAALSAIEMANLFKSLRgsgDYKDDDDKgsgsnmantqmksdkiiiahr gasgylpehtleskalafaqqadyleqdlamtkdgrlvvihdhfldgltdvakkfphrhrkd gryyvidftlkeiqslemtenfetkdgkqaqvypnrfplwkshfrihtfedeiefiqgleks tgkkvgiypeikapwfhhqngkdiaaetlkvlkkygydkktdmvylqtfdfnelkriktell pqmgmdlklvqliaytdwketqekdpkgywvnynydwmfkpgamaevvkyadgvgpgwymlv nkeeskpdnivytplvkelaqynvevhpytvrkdalpefftdvnqmydallnksgatgvftd fpdtgveflkgik
250	TThc-degly-LS	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDgsgdykddddkgsgMQIYEGKLTAEGLRFGIVASRFNHALVDRL VEGCIDCIVRHGGREEDITLVRVPGSWEIPVAAGELARKEDIDAVIAIGVLIRGATPHFDYI ASEVSKGLANLSLELRKPITFGVITADTLEQAIERAGTKHGNKCWEAALSAIEMANLFKSLR gsgDYKDDDDKgsg
251	HID-LS	Snmantqmksdkiiiahrgasgylpehtleskalafaqqadyleqdlamtkdgrlvvihdhf ldgltdvakkfphrhrkdgryyvidftlkeiqslemtenfetkdgkqaqvypnrfplwkshf rihtfedeiefiqglekstgkkvgiypeikapwfhhqngkdiaaetlkvlkkygydkktdmv ylqtfdfnelkriktellpqmgmdlklvqliaytdwketqekdpkgywvnynydwmfkpgam aevvkyadgvgpgwymlvnkeeskpdnivytplvkelaqynvevhpytvrkdalpefftdvn qmydallnksgatgvftdfpdtgveflkgikgsgdykddddkgsgMQIYEGKLTAEGLRFGI VASRFNHALVDRLVEGCIDCIVRHGGREEDITLVRVPGSWEIPVAAGELARKEDIDAVIAIG VLIRGATPHFDYIASEVSKGLANLSLELRKPITFGVITADTLEQAIERAGTKHGNKCWEAAL SAIEMANLFKSLR
252	CRM197-degly-LS	GADDVVDSSKSFVMENFASYHGTKPGYVDSIQKGIQKPKSGTQGNYDDDWKEFYSTDNKYDA AGYSVDNENPLSGKAGGVVKVTYPGLTKVLALKVDNAETIKKELGLSLTEPLMEQVGTEEFI KRFGDGASRVVLSLPFAEGSSSVEYINNWEQAKALSVELEINFETRGKRGQDAMYEYMAQAC AGNRVRRSVGSSLSCINLDWDVIRDKTKTKIESLKEHGPIKNKMSESPNKAVSEEKAKQYLE EFHQTALEHPELSELKTVTGTNPVFAGANYAAWAVNVAQVIDSETADNLEKTTAALSILPGI GSVMGIADGAVHHNTEEIVAQSIALSSLMVAQAIPLVGELVDIGFAAYNFVESIINLFQVVH NSYNRPAYSPGHKTQPFLHDGYAVSWNTVEDSIIRTGFQGESGHDIKITAENTPLPIAGVLL PTIPGKLDVNKAKTHISVNGRKIRMRCRAIDGDVTFCRPKSPVYVGNGVHANLHVAFHRSSS EKIHSNEISSDSIGVLGYQKTVDHTKVNSKLSLFFEIKSgsgdykddddkgsgMQIYEGKLT AEGLRFGIVASRFNHALVDRLVEGCIDCIVRHGGREEDITLVRVPGSWEIPVAAGELARKED IDAVIAIGVLIRGATPHFDYIASEVSKGLANLSLELRKPITFGVITADTLEQAIERAGTKHG NKCWEAALSAIEMANLFKSLR
253	2xCRM197-degly-LS	GADDVVDSSKSFVMENFASYHGTKPGYVDSIQKGIQKPKSGTQGNYDDDWKEFYSTDNKYDA AGYSVDNENPLSGKAGGVVKVTYPGLTKVLALKVDNAETIKKELGLSLTEPLMEQVGTEEFI KRFGDGASRVVLSLPFAEGSSSVEYINNWEQAKALSVELEINFETRGKRGQDAMYEYMAQAC AGNRVRRSVGSSLSCINLDWDVIRDKTKTKIESLKEHGPIKNKMSESPNKAVSEEKAKQYLE EFHQTALEHPELSELKTVTGTNPVFAGANYAAWAVNVAQVIDSETADNLEKTTAALSILPGI GSVMGIADGAVHHNTEEIVAQSIALSSLMVAQAIPLVGELVDIGFAAYNFVESIINLFQVVH NSYNRPAYSPGHKTQPFLHDGYAVSWNTVEDSIIRTGFQGESGHDIKITAENTPLPIAGVLL PTIPGKLDVNKAKTHISVNGRKIRMRCRAIDGDVTFCRPKSPVYVGNGVHANLHVAFHRSSS EKIHSNEISSDSIGVLGYQKTVDHTKVNSKLSLFFEIKSgggsgggsGADDVVDSSKSFVME NFASYHGTKPGYVDSIQKGIQKPKSGTQGNYDDDWKEFYSTDNKYDAAGYSVDNENPLSGKA GGVVKVTYPGLTKVLALKVDNAETIKKELGLSLTEPLMEQVGTEEFIKRFGDGASRVVLSLP FAEGSSSVEYINNWEQAKALSVELEINFETRGKRGQDAMYEYMAQACAGNRVRRSVGSSLSC INLDWDVIRDKTKTKIESLKEHGPIKNKMSESPNKAVSEEKAKQYLEEFHQTALEHPELSEL KTVTGTNPVFAGANYAAWAVNVAQVIDSETADNLEKTTAALSILPGIGSVMGIADGAVHHNT EEIVAQSIALSSLMVAQAIPLVGELVDIGFAAYNFVESIINLFQVVHNSYNRPAYSPGHKTQ

		PFLHDGYAVSWNTVEDSIIRTGFQGESGHDIKITAENTPLPIAGVLLPTIPGKLDVNKAKTH ISVNGRKIRMRCRAIDGDVTFCRPKSPVYVGNGVHANLHVAFHRSSSEKIHSNEISSDSIGV LGYQKTVDHTKVNSKLSLFFEIKSgsgdykdddkgsgMQIYEGKLTAEGLRFGIVASRFNH ALVDRLVEGCIDCIVRHGGREEDITLVRVPGSWEIPVAAGELARKEDIDAVIAIGVLIRGAT PHFDYIASEVSKGLANLSLELRKPITFGVITADTLEQAIERAGTKHGNKCWEAALSAIEMAN LFKSLR
331	LS K7C-R40C caIgG2a-rTT	QIYEGcLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVcHGGREEDITLVRVPGSWEIPVAA GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTLEQAI ERAGTKHGNKGWEAALSAIEMANLFKSLRgsgGSEPKIPQPQPKPQPQPQPQPKPQPKPEPE gsgMKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKA IHLVNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHS LSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANL YINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSY LSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRLY NGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAP GIPLYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHL KDKILGCDWYFVPTDEGWTND
332	LS-L121C-K131C-caIgG2a-rTT	QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAA GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTcEQAI ERAGTcHGNKGWEAALSAIEMANLFKSLRgsgGSEPKIPQPQPKPQPQPQPQPKPQPKPEPE gsgMKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKA IHLVNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHS LSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANL YINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSY LSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRLY NGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAP GIPLYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHL KDKILGCDWYFVPTDEGWTND
333	LS-L121C-K131CG caIgG2a-rTT	QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAA GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTcEQAI ERAGTcgHGNKGWEAALSAIEMANLFKSLRgsgGSEPKIPQPQPKPQPQPQPQPKPQPKPEP EgsgMKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGK AIHLVNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKH SLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSAN LYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTS YLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRL YNGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNA PGIPLYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNH LKDKILGCDWYFVPTDEGWTND
334	LS-L121C-K131GC caIgG2a-rTT	QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAA GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTcEQAI ERAGTgcHGNKGWEAALSAIEMANLFKSLRgsgGSEPKIPQPQPKPQPQPQPQPKPQPKPEP EgsgMKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGK AIHLVNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKH SLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSAN LYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTS YLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRL YNGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNA PGIPLYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNH LKDKILGCDWYFVPTDEGWTND
335	LS-L121CG-K131C caIgG2a-rTT	QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAA GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTcgEQA IERAGTcHGNKGWEAALSAIEMANLFKSLRgsgGSEPKIPQPQPKPQPQPQPQPKPQPKPEP EgsgMKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGK AIHLVNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKH SLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSAN LYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTS YLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRL YNGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNA PGIPLYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNH LKDKILGCDWYFVPTDEGWTND

336	LS-L121GC-K131C caIgG2a-rTT	QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAA GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTgcEQA IERAGTcHGNKGWEAALSAIEMANLFKSLRgsgGSEPKIPQPQPKPQPQPQPQPKPQPKPEP EgsgMKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGK AIHLVNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKH SLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSAN LYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTS YLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRL YNGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNA PGIPLYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNH LKDKILGCDWYFVPTDEGWTND
337	LS K7C-R40C CD8v1-rTT	QIYEGcLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVcHGGREEDITLVRVPGSWEIPVAA GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTLEQAI ERAGTKHGNKGWEAALSAIEMANLFKSLRgsgKPTTTPAPRPPTPAPTIASQPLSLRPEAtR PAAGGAVHTRGgsgMKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPD AQLVPGINGKAIHLVNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNE YSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFIT ITNDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALN PKEIEKLYTSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTN GKLNIYYRRLYNGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNL DRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDI LIASNWYFNHLKDKILGCDWYFVPTDEGWTND
338	LS-L121C-K131C-CD8v1-rTT	QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAA GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTcEQAI ERAGTcHGNKGWEAALSAIEMANLFKSLRgsgKPTTTPAPRPPTPAPTIASQPLSLRPEAtR PAAGGAVHTRGgsgMKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPD AQLVPGINGKAIHLVNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNE YSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFIT ITNDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALN PKEIEKLYTSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTN GKLNIYYRRLYNGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNL DRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDI LIASNWYFNHLKDKILGCDWYFVPTDEGWTND
339	LS-L121C-K131CG CD8v1-rTT	QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAA GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTcEQAI ERAGTcgHGNKGWEAALSAIEMANLFKSLRgsgKPTTTPAPRPPTPAPTIASQPLSLRPEAt RPAAGGAVHTRGgsgMKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYP DAQLVPGINGKAIHLVNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTN EYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFI TITNDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKAL NPKEIEKLYTSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYT NGKLNIYYRRLYNGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNN LDRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRD ILIASNWYFNHLKDKILGCDWYFVPTDEGWTND
340	LS-L121C-K131GC CD8v1-rTT	QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAA GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTcEQAI ERAGTgcHGNKGWEAALSAIEMANLFKSLRgsgKPTTTPAPRPPTPAPTIASQPLSLRPEAt RPAAGGAVHTRGgsgMKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYP DAQLVPGINGKAIHLVNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTN EYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFI TITNDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKAL NPKEIEKLYTSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYT NGKLNIYYRRLYNGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNN LDRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRD ILIASNWYFNHLKDKILGCDWYFVPTDEGWTND
341	LS-L121CG-K131C CD8v1-rTT	QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAA GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTcgEQA IERAGTcHGNKGWEAALSAIEMANLFKSLRgsgKPTTTPAPRPPTPAPTIASQPLSLRPEAt RPAAGGAVHTRGgsgMKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYP DAQLVPGINGKAIHLVNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTN EYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFI TITNDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKAL

		NPKEIEKLYTSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYT NGKLNIYYRRLYNGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNN LDRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRD ILIASNWYFNHLKDKILGCDWYFVPTDEGWTND
342	LS-L121GC-K131C CD8v1-rTT	QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAA GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTgcEQA IERAGTcHGNKGWEAALSAIEMANLFKSLRgsgKPTTTPAPRPPTPAPTIASQPLSLRPEAt RPAAGGAVHTRGgsgMKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYP DAQLVPGINGKAIHLVNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTN EYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFI TITNDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKAL NPKEIEKLYTSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYT NGKLNIYYRRLYNGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNN LDRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRD ILIASNWYFNHLKDKILGCDWYFVPTDEGWTND
343	LS K7C-R40C CD8-rTT	QIYEGcLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVcHGGREEDITLVRVPGSWEIPVAA GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTLEQAI ERAGTKHGNKGWEAALSAIEMANLFKSLRgsgKPTTTPAPRPPTPAPTIASQPLSLRPEACR PAAGGAVHTRGLDFACDgsgMKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSS VITYPDAQLVPGINGKAIHLVNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLE QYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLAN KWVFITITNDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRI FCKALNPKEIEKLYTSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTN APSYTNGKLNIYYRRLYNGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDG NAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGN DPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND
344	LS-L121C-K131C-CD8-rTT	QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAA GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTcEQAI ERAGTcHGNKGWEAALSAIEMANLFKSLRgsgKPTTTPAPRPPTPAPTIASQPLSLRPEACR PAAGGAVHTRGLDFACDgsgMKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSS VITYPDAQLVPGINGKAIHLVNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLE QYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLAN KWVFITITNDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRI FCKALNPKEIEKLYTSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTN APSYTNGKLNIYYRRLYNGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDG NAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGN DPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND
345	LS-L121C-K131CG CD8-rTT	QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAA GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTcEQAI ERAGTcgHGNKGWEAALSAIEMANLFKSLRgsgKPTTTPAPRPPTPAPTIASQPLSLRPEAC RPAAGGAVHTRGLDFACDgsgMKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNS SVITYPDAQLVPGINGKAIHLVNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHL EQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLA NKWVFITITNDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFR IFCKALNPKEIEKLYTSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLT NAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKD GNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIG NDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND
346	LS-L121C-K131GC CD8-rTT	QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAA GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTcEQAI ERAGTgcHGNKGWEAALSAIEMANLFKSLRgsgKPTTTPAPRPPTPAPTIASQPLSLRPEAC RPAAGGAVHTRGLDFACDgsgMKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNS SVITYPDAQLVPGINGKAIHLVNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHL EQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLA NKWVFITITNDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFR IFCKALNPKEIEKLYTSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLT NAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKD GNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIG NDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND
347	LS-L121CG-	QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAA GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTcgEQA IERAGTcHGNKGWEAALSAIEMANLFKSLRgsgKPTTTPAPRPPTPAPTIASQPLSLRPEAC

	K131C CD8-rTT	RPAAGGAVHTRGLDFACDgsgMKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNS SVITYPDAQLVPGINGKAIHLVNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHL EQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLA NKWVFITITNDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFR IFCKALNPKEIEKLYTSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLT NAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKD GNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIG NDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND
348	LS-L121GC-K131C CD8-rTT	QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAA GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTgcEQA IERAGTcHGNKGWEAALSAIEMANLFKSLRgsgKPTTTPAPRPPTPAPTIASQPLSLRPEAC RPAAGGAVHTRGLDFACDgsgMKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNS SVITYPDAQLVPGINGKAIHLVNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHL EQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLA NKWVFITITNDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFR IFCKALNPKEIEKLYTSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLT NAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKD GNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIG NDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND
349	LS K7C-R40C hinge-rTT	QIYEGcLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVcHGGREEDITLVRVPGSWEIPVAA GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTLEQAI ERAGTKHGNKGWEAALSAIEMANLFKSLRggsgEPKSDKTHTPPPAPELLgsgEPKSDKTHT PPPAPELLgsggMKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQ LVPGINGKAIHLVNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYS IISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITIT NDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPK EIEKLYTSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGK LNIYYRRLYNGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDR ILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILI ASNWYFNHLKDKILGCDWYFVPTDEGWTND
350	LS-L121C-K131C hinge-rTT	QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAA GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTcEQAI ERAGTcHGNKGWEAALSAIEMANLFKSLRggsgEPKSDKTHTPPPAPELLgsgEPKSDKTHT PPPAPELLgsggMKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQ LVPGINGKAIHLVNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYS IISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITIT NDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPK EIEKLYTSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGK LNIYYRRLYNGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDR ILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILI ASNWYFNHLKDKILGCDWYFVPTDEGWTND
351	LS-L121C-K131CG hinge-rTT	QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAA GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTcEQAI ERAGTcgHGNKGWEAALSAIEMANLFKSLRggsgEPKSDKTHTPPPAPELLgsgEPKSDKTH TPPPAPELLgsggMKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDA QLVPGINGKAIHLVNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEY SIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITI TNDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNP KEIEKLYTSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNG KLNIYYRRLYNGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLD RILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDIL IASNWYFNHLKDKILGCDWYFVPTDEGWTND
352	LS-L121C-K131GC hinge-rTT	QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAA GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTcEQAI ERAGTgcHGNKGWEAALSAIEMANLFKSLRggsgEPKSDKTHTPPPAPELLgsgEPKSDKTH TPPPAPELLgsggMKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDA QLVPGINGKAIHLVNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEY SIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITI TNDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNP KEIEKLYTSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNG KLNIYYRRLYNGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLD

		RILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDIL IASNWYFNHLKDKILGCDWYFVPTDEGWTND
353	LS-L121CG-K131C hinge-rTT	QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAA GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTcgEQA IERAGTcHGNKGWEAALSAIEMANLFKSLRggsgEPKSDKTHTPPPAPELLgsgEPKSDKTH TPPPAPELLgsggMKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDA QLVPGINGKAIHLVNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEY SIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITI TNDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNP KEIEKLYTSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNG KLNIYYRRLYNGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLD RILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDIL IASNWYFNHLKDKILGCDWYFVPTDEGWTND
354	LS-L121GC-K131C hinge-rTT	QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAA GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTgcEQA IERAGTcHGNKGWEAALSAIEMANLFKSLRggsgEPKSDKTHTPPPAPELLgsgEPKSDKTH TPPPAPELLgsggMKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDA QLVPGINGKAIHLVNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEY SIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITI TNDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNP KEIEKLYTSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNG KLNIYYRRLYNGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLD RILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDIL IASNWYFNHLKDKILGCDWYFVPTDEGWTND
355	LS-15-TT	QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAA GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTLEQAI ERAGTKHGNKGWEAALSAIEMANLFKSLRggggsggggsggggsMKNLDCWVDNEEDIDVIL KKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAIHLVNNESSEVIVHKAMDIEY NDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLK DSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGVLMGSAEITGLGAIRED NNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITFLRDFWGNPLRYDTEYY LIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNEIDSFVK SGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDLKTYSV QLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND
356	LS-25-TT	QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAA GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTLEQAI ERAGTKHGNKGWEAALSAIEMANLFKSLRggksggnkksdgvkessesgggsggMKNLDCWV DNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAIHLVNNESSEV IVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSL KGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGVLMGSAE ITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITFLRDFWG NPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYT PNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAV KLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYF VPTDEGWTND
357	LS-30-TT	QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAA GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTLEQAI ERAGTKHGNKGWEAALSAIEMANLFKSLRggggsggksggnkksdgvkessesgggsggMKN LDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAIHLVNN ESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSG WSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGVL MGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITFL RDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFI IKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYK KMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILG CDWYFVPTDEGWTND
358	LS-35-TT	QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAA GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTLEQAI ERAGTKHGNKGWEAALSAIEMANLFKSLRggggsggksggnkksdgvkessesgggsggggg gsMKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAI RLVNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSL SIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLY

		INGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYL SITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRLYN GLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPG IPLYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLK DKILGCDWYFVPTDEGWTND
359	LS-20-env31-TT	QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAA GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTLEQAI ERAGTKHGNKGWEAALSAIEMANLFKSLRggksggnkksdgvkSLVRAENLWVTVYYGVPVW klvtgessesgMKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQL VPGINGKAIHLVNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSI ISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITN DRLSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKE IEKLYTSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKL NIYYRRLYNGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRI LRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIA SNWYFNHLKDKILGCDWYFVPTDEGWTND
360	LS-20-PADRE-TT	QIYEGKLTAEGLRFGIVASPFNHALVDRLVEGAIDAIVPHGGREEDITLVRVPGSWEIPVAA GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELPKPITFGVITADTLEQAI ERAGTRHGNKGWEAALSAIEMANLFKSLKggksggnkksdgvkeSLVRAKFVAAWTLKAAAG SLVRssesgMKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVP GINGKAIHLVNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIIS SMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDR LSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIE KLYTSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNI YYRRLYNGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILR VGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASN WYFNHLKDKILGCDWYFVPTDEGWTND
361	LS-hinge-rTT	QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAA GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTLEQAI ERAGTKHGNKGWEAALSAIEMANLFKSLRggEPKSCDKTHTCPPCPAPELLggKNLDCWVDN EEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAIHLVNNESSEVIV HKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKG NNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGVLMGSAEIT GLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITFLRDFWGNP LRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPN NEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKL RDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFVP TDEGWTND
362	LS-hinge2-rTT (remove the cys from hinge)	QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAA GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTLEQAI ERAGTKHGNKGWEAALSAIEMANLFKSLRggEPKSDKTHTPPPAPELLggKNLDCWVDNEED IDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAIHLVNNESSEVIVHKA MDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNL IWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGVLMGSAEITGLG AIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITFLRDFWGNPLRY DTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNEI DSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDL KTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFVPTDE GWTND
363	LS-hinge2.1-rTT	QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAA GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTLEQAI ERAGTKHGNKGWEAALSAIEMANLFKSLRggsgEPKSDKTHTPPPAPELLgsgEPKSDKTHT PPPAPELLgsggMKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQ LVPGINGKAIHLVNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYS IISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITIT NDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPK EIEKLYTSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGK LNIYYRRLYNGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDR ILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILI ASNWYFNHLKDKILGCDWYFVPTDEGWTND
364	LS-hinge3-	QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAA GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTLEQAI

	rTT (mutate the cys to Thr in hinge)	ERAGTKHGNKGWEAALSAIEMANLFKSLRggEPKSTDKTHTSPPSPAPELLggKNLDCWVDN EEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAIHLVNNESSEVIV HKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKG NNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGVLMGSAEIT GLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITFLRDFWGNP LRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPN NEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKL RDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFVP TDEGWTND
365	LS-ext1-rTT (extend the N terminal of rTT)	QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAA GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTLEQAI ERAGTKHGNKGWEAALSAIEMANLFKSLRggGITELKKLESKINKVFSTPIPFSYSKNLDCWV DNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAIHLVNNESSEV IVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSL KGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGVLMGSAE ITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITFLRDFWG NPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYT PNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAV KLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYF VPTDEGWTND
366	LS-caIgG2a-rTT	QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAA GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTLEQAI ERAGTKHGNKGWEAALSAIEMANLFKSLRgsgGSEPKIPQPQPKPQPQPQPQPKPQPKPEPE gsgMKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKA IHLVNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHS LSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANL YINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSY LSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRLY NGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAP GIPLYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHL KDKILGCDWYFVPTDEGWTND
367	LS-CD8-rTT	QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAA GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTLEQAI ERAGTKHGNKGWEAALSAIEMANLFKSLRgsgKPTTTPAPRPPTPAPTIASQPLSLRPEAtR PAAGGAVHTRGgsgMKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPD AQLVPGINGKAIHLVNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNE YSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFIT ITNDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALN PKEIEKLYTSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTN GKLNIYYRRLYNGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNL DRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDI LIASNWYFNHLKDKILGCDWYFVPTDEGWTND
368	LS-CD8v2-rTT	QIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVAA GELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTLEQAI ERAGTKHGNKGWEAALSAIEMANLFKSLRgsgKPTTTPAPRPPTPAPTIASQPLSLRPEACR PAAGGAVHTRGLDFACDgsgMKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSS VITYPDAQLVPGINGKAIHLVNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLE QYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLAN KWVFITITNDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRI FCKALNPKEIEKLYTSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTN APSYTNGKLNIYYRRLYNGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDG NAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGN DPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND
Ferritin		
109	rTT-ferr1	NLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNS**A**VITYPDAQLVPGINGKAIHLVN NE**A**SEVIVHKAMDIEYNDMFN**Q**FTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGS GWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGV LMGSAEITGLGAIREDN**Q**ITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITF LRDFWGNPLRYDTEYYLIPVASSSKDVQLK**Q**ITDYMYLTNAPSYTNGKLNIYYRRLYNGLKF IIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLY KKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKIL

		GCDWYFVPTDEGWTNDggsgggsgg*ASISEKMVEALNRQINAEIYSAYLYLSMASYFDSIGL KGFSNWMRVQWQEELMHAMKMFDFVSRRGGRVKLYAVEEPPSEWDSPLAAFEHVYEHEVNVV KRIHELVEMAMQEKDFATYNFLQWYVAEQVEEEASALDIVEKLRLIGEDKRALLFLDKELSL RQFTPPAEEEK*
110	rTT-ferr 2	NLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNS**A**VITYPDAQLVPGINGKAIHLVN NE**A**SEVIVHKAMDIEYNDMFN**Q**FTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGS GWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGV LMGSAEITGLGAIREDN**Q**ITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITF LRDFWGNPLRYDTEYYLIPVASSSKDVQLK**Q**ITDYMYLTNAPSYTNGKLNIYYRRLYNGLKF IIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLY KKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKIL GCDWYFVPTDEGWTNDggsgggsgggsggsg*ASISEKMVEALNRQINAEIYSAYLYLSMASY FDSIGLKGFSNWMRVQWQEELMHAMKMFDFVSRRGGRVKLYAVEEPPSEWDSPLAAFEHVYE HEVNVVKRIHELVEMAMQEKDFATYNFLQWYVAEQVEEEASALDIVEKLRLIGEDKRALLFL DKELSLRQFTPPAEEEK*
111	rTT-ferr 3	NLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNS**A**VITYPDAQLVPGINGKAIHLVN NE**A**SEVIVHKAMDIEYNDMFN**Q**FTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGS GWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGV LMGSAEITGLGAIREDN**Q**ITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITF LRDFWGNPLRYDTEYYLIPVASSSKDVQLK**Q**ITDYMYLTNAPSYTNGKLNIYYRRLYNGLKF IIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLY KKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKIL GCDWYFVPTDEGWTNDggdgggdgggdgggdg*ASISEKMVEALNRQINAEIYSAYLYLSMASY FDSIGLKGFSNWMRVQWQEELMHAMKMFDFVSRRGGRVKLYAVEEPPSEWDSPLAAFEHVYE HEVNVVKRIHELVEMAMQEKDFATYNFLQWYVAEQVEEEASALDIVEKLRLIGEDKRALLFL DKELSLRQFTPPAEEEK*
112	rTT-ferr 4	NLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNS**A**VITYPDAQLVPGINGKAIHLVN NE**A**SEVIVHKAMDIEYNDMFN**Q**FTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGS GWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGV LMGSAEITGLGAIREDN**Q**ITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITF LRDFWGNPLRYDTEYYLIPVASSSKDVQLK**Q**ITDYMYLTNAPSYTNGKLNIYYRRLYNGLKF IIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLY KKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKIL GCDWYFVPTDEGWTNDggsgggsgg*MLSKDIIKLLNEQVNKEMDSSNLYMSMSSWCYTHSLD GAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKFEGLTQIFQKAYEHEQHISE SINNIVDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDKIELIGNENHGLYLADQYVKGI AKSRKS*
113	rTT-ferr 5	NLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNS**A**VITYPDAQLVPGINGKAIHLVN NE**A**SEVIVHKAMDIEYNDMFN**Q**FTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGS GWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGV LMGSAEITGLGAIREDN**Q**ITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITF LRDFWGNPLRYDTEYYLIPVASSSKDVQLK**Q**ITDYMYLTNAPSYTNGKLNIYYRRLYNGLKF IIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLY KKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKIL GCDWYFVPTDEGWTNDggsgggsgggsggsg*MLSKDIIKLLNEQVNKEMDSSNLYMSMSSWC YTHSLDGAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKFEGLTQIFQKAYEH EQHISESINNIVDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDKIELIGNENHGLYLAD QYVKGIAKSRKS*
114	rTT-ferr 6	NLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNS**A**VITYPDAQLVPGINGKAIHLVN NE**A**SEVIVHKAMDIEYNDMFN**Q**FTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGS GWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGV LMGSAEITGLGAIREDN**Q**ITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITF LRDFWGNPLRYDTEYYLIPVASSSKDVQLK**Q**ITDYMYLTNAPSYTNGKLNIYYRRLYNGLKF IIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLY KKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKIL GCDWYFVPTDEGWTNDggdgggdgggdgggdg*MLSKDIIKLLNEQVNKEMDSSNLYMSMSSWC YTHSLDGAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKFEGLTQIFQKAYEH EQHISESINNIVDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDKIELIGNENHGLYLAD QYVKGIAKSRKS*
115	rTT_degly -Fer	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN

		GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDgggSGG*DIIKLLNEQVNKEMQSSNLYMSMSSWCYTHSLDGAGLF* *LFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKFEGLTQIFQKAYEHEQHISESINNI* *VDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDKIELIGNENHGLYLADQYVKGIAKSRK* *S*
116	rTT_degly -ln15-Fer	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDgggSGGggsggSGGg*DIIKLLNEQVNKEMQSSNLYMSMSSWCYT* *HSLDGAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKFEGLTQIFQKAYEHEQ* *HISESINNIVDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDKIELIGNENHGLYLADQY* *VKGIAKSRKS*
117	rTT_degly -ln25-Fer	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDgggSGGggsggSGGggSGGggSGGg*DIIKLLNEQVNKEMQSSNL* *YMSMSSWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKFEGLTQ* *IFQKAYEHEQHISESINNIVDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDKIELIGNE* *NHGLYLADQYVKGIAKSRKS*
118	rTT_degly -ln35-Fer	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDgggSGGggsggSGGggSGGggSGGggSGGggSGGg*DIIKLLNEQ* *VNKEMQSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQLTSISA* *PEHKFEGLTQIFQKAYEHEQHISESINNIVDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDI* *LDKIELIGNENHGLYLADQYVKGIAKSRKS*
119	rTT_degly -K5A-Fer	MKNLDCWVDkEEDIDVILKKSTILNLDINkDIISDISkFNSAVITYPDAQLVPGINGKAIHL VNNEkSEVIVHKAMkIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDgggSGG*DIIKLLNEQVNKEMQSSNLYMSMSSWCYTHSLDGAGLF* *LFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKFEGLTQIFQKAYEHEQHISESINNI* *VDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDKIELIGNENHGLYLADQYVKGIAKSRK* *S*
120	rTT_degly -K5B-Fer	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHkkSI GSGWSVSLKGNNLIWTLKDSkGEVRQITFRDLPkKFNAYLANKWVFITITNDRkSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDgggSGG*DIIKLLNEQVNKEMQSSNLYMSMSSWCYTHSLDGAGLF* *LFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKFEGLTQIFQKAYEHEQHISESINNI*

		VDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDKIELIGNENHGLYLADQYVKGIAKSRK S
121	rTT_degly -K5C-Fer	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITkLGAIREDNQITLKLDRCkNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTkAPSYTNGKLNIYYRRLYNGL KFIIKRYkPNNkIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDgggSGG*DIIKLLNEQVNKEMQSSNLYMSMSSWCYTHSLDGAGLF LFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKFEGLTQIFQKAYEHEQHISESINNI VDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDKIELIGNENHGLYLADQYVKGIAKSRK S*
122	rTT_degly -K5D-Fer	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYkNNEHIVGYPKDGNAFNkLDRILRVGYkAPkIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGkDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDgggSGG*DIIKLLNEQVNKEMQSSNLYMSMSSWCYTHSLDGAGLF LFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKFEGLTQIFQKAYEHEQHISESINNI VDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDKIELIGNENHGLYLADQYVKGIAKSRK S*
123	rTT_degly -K10A-Fer	MKNLDCWVDkEEDIDVILKKSTILNLDINkDIISDISkFNSAVITYPDAQLVPGINGKAIHL VNNEkSEVIVHKAMkIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHkkSI GSGWSVSLKGNNLIWTLKDSkGEVRQITFRDLPkKFNAYLANKWVFITITNDRkSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDgggSGG*DIIKLLNEQVNKEMQSSNLYMSMSSWCYTHSLDGAGLF LFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKFEGLTQIFQKAYEHEQHISESINNI VDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDKIELIGNENHGLYLADQYVKGIAKSRK S*
124	rTT_degly -K10B-Fer	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITkLGAIREDNQITLKLDRCkNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTkAPSYTNGKLNIYYRRLYNGL KFIIKRYkPNNkIDSFVKSGDFIKLYVSYkNNEHIVGYPKDGNAFNkLDRILRVGYkAPkIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGkDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDgggSGG*DIIKLLNEQVNKEMQSSNLYMSMSSWCYTHSLDGAGLF LFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKFEGLTQIFQKAYEHEQHISESINNI VDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDKIELIGNENHGLYLADQYVKGIAKSRK S*
125	rTT_degly -K10C-Fer	MKNLDCWVDkEEDIDVILKKSTILNLDINkDIISDISkFNSAVITYPDAQLVPGINGKAIHL VNNEkSEVIVHKAMkIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITkLGAIREDNQITLKLDRCkNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTkAPSYTNGKLNIYYRRLYNGL KFIIKRYkPNNkIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDgggSGG*DIIKLLNEQVNKEMQSSNLYMSMSSWCYTHSLDGAGLF LFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKFEGLTQIFQKAYEHEQHISESINNI VDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDKIELIGNENHGLYLADQYVKGIAKSRK S*
126	rTT_degly -K10D-Fer	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHkkSI GSGWSVSLKGNNLIWTLKDSkGEVRQITFRDLPkKFNAYLANKWVFITITNDRkSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL

		KFIIKRYTPNNEIDSFVKSGDFIKLYVSYkNNEHIVGYPKDGNAFNkLDRILRVGYkAPkIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGkDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDgggSGG*DIIKLLNEQVNKEMQSSNLYMSMSSWCYTHSLDGAGLF* *LFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKFEGLTQIFQKAYEHEQHISESINNI* *VDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDKIELIGNENHGLYLADQYVKGIAKSRK* *S*
127	rTT_degly -K15A-Fer	MKNLDCWVDkEEDIDVILKKSTILNLDINkDIISDISkFNSAVITYPDAQLVPGINGKAIHL VNNEkSEVIVHKAMkIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHkkSI GSGWSVSLKGNNLIWTLKDSkGEVRQITFRDLPkKFNAYLANKWVFITITNDRkSSANLYIN GVLMGSAEITkLGAIREDNQITLKLDRCkNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTkAPSYTNGKLNIYYRRLYNGL KFIIKRYkPNNkIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDgggSGG*DIIKLLNEQVNKEMQSSNLYMSMSSWCYTHSLDGAGLF* *LFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKFEGLTQIFQKAYEHEQHISESINNI* *VDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDKIELIGNENHGLYLADQYVKGIAKSRK* *S*
128	rTT_degly -K15B-Fer	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHkkSI GSGWSVSLKGNNLIWTLKDSkGEVRQITFRDLPkKFNAYLANKWVFITITNDRkSSANLYIN GVLMGSAEITkLGAIREDNQITLKLDRCkNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTkAPSYTNGKLNIYYRRLYNGL KFIIKRYkPNNkIDSFVKSGDFIKLYVSYkNNEHIVGYPKDGNAFNkLDRILRVGYkAPkIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGkDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDgggSGG*DIIKLLNEQVNKEMQSSNLYMSMSSWCYTHSLDGAGLF* *LFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKFEGLTQIFQKAYEHEQHISESINNI* *VDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDKIELIGNENHGLYLADQYVKGIAKSRK* *S*
129	rTT_degly -K15C-Fer	MKNLDCWVDkEEDIDVILKKSTILNLDINkDIISDISkFNSAVITYPDAQLVPGINGKAIHL VNNEkSEVIVHKAMkIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITkLGAIREDNQITLKLDRCkNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTkAPSYTNGKLNIYYRRLYNGL KFIIKRYkPNNkIDSFVKSGDFIKLYVSYkNNEHIVGYPKDGNAFNkLDRILRVGYkAPkIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGkDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDgggSGG*DIIKLLNEQVNKEMQSSNLYMSMSSWCYTHSLDGAGLF* *LFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKFEGLTQIFQKAYEHEQHISESINNI* *VDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDKIELIGNENHGLYLADQYVKGIAKSRK* *S*
130	rTT_degly -K15D-Fer	MKNLDCWVDkEEDIDVILKKSTILNLDINkDIISDISkFNSAVITYPDAQLVPGINGKAIHL VNNEkSEVIVHKAMkIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHkkSI GSGWSVSLKGNNLIWTLKDSkGEVRQITFRDLPkKFNAYLANKWVFITITNDRkSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYkNNEHIVGYPKDGNAFNkLDRILRVGYkAPkIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGkDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDgggSGG*DIIKLLNEQVNKEMQSSNLYMSMSSWCYTHSLDGAGLF* *LFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKFEGLTQIFQKAYEHEQHISESINNI* *VDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDKIELIGNENHGLYLADQYVKGIAKSRK* *S*
131	rTT_degly -K20-Fer	MKNLDCWVDkEEDIDVILKKSTILNLDINkDIISDISkFNSAVITYPDAQLVPGINGKAIHL VNNEkSEVIVHKAMkIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHkkSI GSGWSVSLKGNNLIWTLKDSkGEVRQITFRDLPkKFNAYLANKWVFITITNDRkSSANLYIN GVLMGSAEITkLGAIREDNQITLKLDRCkNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTkAPSYTNGKLNIYYRRLYNGL KFIIKRYkPNNkIDSFVKSGDFIKLYVSYkNNEHIVGYPKDGNAFNkLDRILRVGYkAPkIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGkDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDgggSGG*DIIKLLNEQVNKEMQSSNLYMSMSSWCYTHSLDGAGLF* *LFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKFEGLTQIFQKAYEHEQHISESINNI* *VDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDKIELIGNENHGLYLADQYVKGIAKSRK* *S*

132	rTT_degly-K20-ln15-Fer	MKNLDCWVDkEEDIDVILKKSTILNLDINkDIISDISkFNSAVITYPDAQLVPGINGKAIHL VNNEkSEVIVHKAMkIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHkkSI GSGWSVSLKGNNLIWTLKDSkGEVRQITFRDLPkKFNAYLANKWVFITITNDRkSSANLYIN GVLMGSAEITkLGAIREDNQITLKLDRCkNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTkAPSYTNGKLNIYYRRLYNGL KFIIKRYkPNNkIDSFVKSGDFIKLYVSYkNNEHIVGYPKDGNAFNkLDRILRVGYkAPkIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGkDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDgggSGGggsggSGGg*DIIKLLNEQVNKEMQSSNLYMSMSSWCYT HSLDGAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKFEGLTQIFQKAYEHEQ HISESINNIVDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDKIELIGNENHGLYLADQY VKGIAKSRKS*
133	rTT_degly-K20-ln25-Fer	MKNLDCWVDkEEDIDVILKKSTILNLDINkDIISDISkFNSAVITYPDAQLVPGINGKAIHL VNNEkSEVIVHKAMkIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHkkSI GSGWSVSLKGNNLIWTLKDSkGEVRQITFRDLPkKFNAYLANKWVFITITNDRkSSANLYIN GVLMGSAEITkLGAIREDNQITLKLDRCkNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTkAPSYTNGKLNIYYRRLYNGL KFIIKRYkPNNkIDSFVKSGDFIKLYVSYkNNEHIVGYPKDGNAFNkLDRILRVGYkAPkIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGkDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDgggSGGggsggSGGggSGGggSGGg*DIIKLLNEQVNKEMQSSNL YMSMSSWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKFEGLTQ IFQKAYEHEQHISESINNIVDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDKIELIGNE NHGLYLADQYVKGIAKSRKS*
134	rTT_degly-K20-ln35-Fer	MKNLDCWVDkEEDIDVILKKSTILNLDINkDIISDISkFNSAVITYPDAQLVPGINGKAIHL VNNEkSEVIVHKAMkIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHkkSI GSGWSVSLKGNNLIWTLKDSkGEVRQITFRDLPkKFNAYLANKWVFITITNDRkSSANLYIN GVLMGSAEITkLGAIREDNQITLKLDRCkNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTkAPSYTNGKLNIYYRRLYNGL KFIIKRYkPNNkIDSFVKSGDFIKLYVSYkNNEHIVGYPKDGNAFNkLDRILRVGYkAPkIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGkDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDgggSGGggsggSGGggSGGggSGGggSGGggSGGg*DIIKLLNEQ VNKEMQSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQLTSISA PEHKFEGLTQIFQKAYEHEQHISESINNIVDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDI LDKIELIGNENHGLYLADQYVKGIAKSRKS*
135	rTT_degly-r8-Ferr	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDggaeaaakeaaakeaaakeaaakaleaeaaakeaaakeaaakea aaka*DIIKLLNEQVNKEMQSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIFLN ENNVPVQLTSISAPEHKFEGLTQIFQKAYEHEQHISESINNIVDHAIKSKDHATFNFLQWYV AEQHEEEVLFKDILDKIELIGNENHGLYLADQYVKGIAKSRKS*
136	rTT_degly-12pa-Ferr	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDgsapapapapapapapapapapapapa*DIIKLLNEQVNKEMQSS NLYMSMSSWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKFEGL TQIFQKAYEHEQHISESINNIVDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDKIELIG NENHGLYLADQYVKGIAKSRKS*
137	rTT_degly-r3-Ferr	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK

		ILGCDWYFVPTDEGWTNDggaeaaakeaaakeaaaka*DIIKLLNEQVNKEMQSSNLYMSMSS WCYTHSLDGAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKFEGLTQIFQKAY EHEQHISESINNIVDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDKIELIGNENHGLYL ADQYVKGIAKSRKS*
138	rTT_degly -3f-Ferr	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDggggsggggsggggs*DIIKLLNEQVNKEMQSSNLYMSMSSWCYT HSLDGAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKFEGLTQIFQKAYEHEQ HISESINNIVDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDKIELIGNENHGLYLADQY VKGIAKSRKS*
139	rTT_degly -2rf-Ferr	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDgsaeaaakeaaaka*DIIKLLNEQVNKEMQSSNLYMSMSSWCYTH SLDGAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKFEGLTQIFQKAYEHEQH ISESINNIVDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDKIELIGNENHGLYLADQYV KGIAKSRKS*
140	rTT_degly -5ga-Ferr	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDgsgsgsgsgsgsasg*DIIKLLNEQVNKEMQSSNLYMSMSSWCYTHS LDGAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKFEGLTQIFQKAYEHEQHI SESINNIVDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDKIELIGNENHGLYLADQYVK GIAKSRKS*
141	rTT_degly -1rf-Ferr	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDpapapasg*DIIKLLNEQVNKEMQSSNLYMSMSSWCYTHSLDGAG LFLFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKFEGLTQIFQKAYEHEQHISESIN NIVDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDKIELIGNENHGLYLADQYVKGIAKS RKS*
142	Ferr_delt aCT1-linker-rTT	*DIIKLLNEQVNKEMNSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIFLNENNV PVQLTSISAPEHKFEGLTQIFQKAYEHEQHISESINNIVDHAIKSKDHATFNFLQWYVAEQH EEEVLFKDILDKIELIG***KALEAQKQK***MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDI* SGFNSSVITYPDAQLVPGINGKAIHLVNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKV SASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKF NAYLANKWVFITITNDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVS IDKFRIFCKALNPKEIEKLYTSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITD YMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIV GYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTH NGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND
143	Ferr_delt aCT2-linker-rTT	*DIIKLLNEQVNKEMNSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIFLNENNV PVQLTSISAPEHKFEGLTQIFQKAYEHEQHISESINNIVDHAIKSKDHATFNFLQWYVAEQH EEEVLFKDILDKIELIGNEN***KALEAQKQK***MKNLDCWVDNEEDIDVILKKSTILNLDINNDII* SDISGFNSSVITYPDAQLVPGINGKAIHLVNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRV

		PKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLP DKFNAYLANKWVFITITNDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQ YVSIDKFRIFCKALNPKEIEKLYTSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKN ITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNE HIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLV GTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND
144	rTT-linker-Ferr	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAIHL VNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTND**KALEAQKQK***SKDIIKLLNEQVNKEMNSSNLYMSMSSWCYTHSLD GAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKFEGLTQIFQKAYEHEQHISE SINNIVDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDKIELIGNENHGLYLADQYVKGI AKSRKS*
145	rTT-2xlinker-Ferr	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAIHL VNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTND**KALEAQKQKKALEAQKQK***SKDIIKLLNEQVNKEMNSSNLYMSMS SWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEHKFEGLTQIFQKA YEHEQHISESINNIVDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDKIELIGNENHGLY LADQYVKGIAKSRKS*
146	Ferr_deltaCT1-linker-CRM	*DIIKLLNEQVNKEMNSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIFLNENNV PVQLTSISAPEHKFEGLTQIFQKAYEHEQHISESINNIVDHAIKSKDHATFNFLQWYVAEQH EEEVLFKDILDKIELIG**KALEAQKQK**GADDVVDSSKSFVMENFSSYHGTKPGYVDSIQKGIQ KPKSGTQGNYDDDWKEFYSTDNKYDAAGYSVDNENPLSGKAGGVVKVTYPGLTKVLALKVDN AETIKKELGLSLTEPLMEQVGTEEFIKRFGDGASRVVLSLPFAEGSSSVEYINNWEQAKALS VELEINFETRGKRGQDAMYEYMAQACAGNRVRRSVGSSLSCINLDWDVIRDKTKTKIESLKE HGPIKNKMSESPNKTVSEEKAKQYLEEFHQTALEHPELSELKTVTGTNPVFAGANYAAWAVN VAQVIDSETADNLEKTTAALSILPGIGSVMGIADGAVHHNTEEIVAQSIALSSLMVAQAIPL VGELVDIGFAAYNFVESIINLFQVVHNSYNRPAYSPGHKTQPFLHDGYAVSWNTVEDSIIRT GFQGESGHDIKITAENTPLPIAGVLLPTIPGKLDVNKSKTHISVNGRKIRMRCRAIDGDVTF CRPKSPVYVGNGVHANLHVAFHRSSSEKIHSNEISSDSIGVLGYQKTVDHTKVNSKLSLFFE IKS*
147	Ferr_deltaCT2-linker-CRM	*DIIKLLNEQVNKEMNSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIFLNENNV PVQLTSISAPEHKFEGLTQIFQKAYEHEQHISESINNIVDHAIKSKDHATFNFLQWYVAEQH EEEVLFKDILDKIELIGNEN**KALEAQKQK**GADDVVDSSKSFVMENFSSYHGTKPGYVDSIQK GIQKPKSGTQGNYDDDWKEFYSTDNKYDAAGYSVDNENPLSGKAGGVVKVTYPGLTKVLALK VDNAETIKKELGLSLTEPLMEQVGTEEFIKRFGDGASRVVLSLPFAEGSSSVEYINNWEQAK ALSVELEINFETRGKRGQDAMYEYMAQACAGNRVRRSVGSSLSCINLDWDVIRDKTKTKIES LKEHGPIKNKMSESPNKTVSEEKAKQYLEEFHQTALEHPELSELKTVTGTNPVFAGANYAAW AVNVAQVIDSETADNLEKTTAALSILPGIGSVMGIADGAVHHNTEEIVAQSIALSSLMVAQA IPLVGELVDIGFAAYNFVESIINLFQVVHNSYNRPAYSPGHKTQPFLHDGYAVSWNTVEDSI IRTGFQGESGHDIKITAENTPLPIAGVLLPTIPGKLDVNKSKTHISVNGRKIRMRCRAIDGD VTFCRPKSPVYVGNGVHANLHVAFHRSSSEKIHSNEISSDSIGVLGYQKTVDHTKVNSKLSL FFEIKS*
148	CRM-linker-Ferr	GADDVVDSSKSFVMENFSSYHGTKPGYVDSIQKGIQKPKSGTQGNYDDDWKEFYSTDNKYDA AGYSVDNENPLSGKAGGVVKVTYPGLTKVLALKVDNAETIKKELGLSLTEPLMEQVGTEEFI KRFGDGASRVVLSLPFAEGSSSVEYINNWEQAKALSVELEINFETRGKRGQDAMYEYMAQAC AGNRVRRSVGSSLSCINLDWDVIRDKTKTKIESLKEHGPIKNKMSESPNKTVSEEKAKQYLE EFHQTALEHPELSELKTVTGTNPVFAGANYAAWAVNVAQVIDSETADNLEKTTAALSILPGI GSVMGIADGAVHHNTEEIVAQSIALSSLMVAQAIPLVGELVDIGFAAYNFVESIINLFQVVH NSYNRPAYSPGHKTQPFLHDGYAVSWNTVEDSIIRTGFQGESGHDIKITAENTPLPIAGVLL PTIPGKLDVNKSKTHISVNGRKIRMRCRAIDGDVTFCRPKSPVYVGNGVHANLHVAFHRSSS EKIHSNEISSDSIGVLGYQKTVDHTKVNSKLSLFFEIKS**KALEAQKQK***SKDIIKLLNEQVNK EMNSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQLTSISAPEH

		KFEGLTQIFQKAYEHEQHISESINNIVDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDK IELIGNENHGLYLADQYVKGIAKSRKS
149	CRM-2xlinker-Ferr	GADDVVDSSKSFVMENFSSYHGTKPGYVDSIQKGIQKPKSGTQGNYDDDWKEFYSTDNKYDA AGYSVDNENPLSGKAGGVVKVTYPGLTKVLALKVDNAETIKKELGLSLTEPLMEQVGTEEFI KRFGDGASRVVLSLPFAEGSSSVEYINNWEQAKALSVELEINFETRGKRGQDAMYEYMAQAC AGNRVRRSVGSSLSCINLDWDVIRDKTKTKIESLKEHGPIKNKMSESPNKTVSEEKAKQYLE EFHQTALEHPELSELKTVTGTNPVFAGANYAAWAVNVAQVIDSETADNLEKTTAALSILPGI GSVMGIADGAVHHNTEEIVAQSIALSSLMVAQAIPLVGELVDIGFAAYNFVESIINLFQVVH NSYNRPAYSPGHKTQPFLHDGYAVSWNTVEDSIIRTGFQGESGHDIKITAENTPLPIAGVLL PTIPGKLDVNKSKTHISVNGRKIRMRCRAIDGDVTFCRPKSPVYVGNGVHANLHVAFHRSSS EKIHSNEISSDSIGVLGYQKTVDHTKVNSKLSLFFEIKS**KALEAQKQKKALEAQKQK***SKDII KLLNEQVNKEMNSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQ LTSISAPEHKFEGLTQIFQKAYEHEQHISESINNIVDHAIKSKDHATFNFLQWYVAEQHEEE VLFKDILDKIELIGNENHGLYLADQYVKGIAKSRKS*
150	rTT-ferr 1	NLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNS**A**VITYPDAQLVPGINGKAIHLVN NE**A**SEVIVHKAMDIEYNDMFN**Q**FTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGS GWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGV LMGSAEITGLGAIREDN**Q**ITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITF LRDFWGNPLRYDTEYYLIPVASSSKDVQLK**Q**ITDYMYLTNAPSYTNGKLNIYYRRLYNGLKF IIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLY KKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKIL GCDWYFVPTDEGWTND*ggsggggsgg**ASISEKMVEALNRQINAEIYSAYLYLSMASYFDSIGL KGFSNWMRVQWQEELMHAMKMFDFVSRRGGRVKLYAVEEPPSEWDSPLAAFEHVYEHEVNVV KRIHELVEMAMQEKDFATYNFLQWYVAEQVEEEASALDIVEKLRLIGEDKRALLFLDKELSL RQFTPPAEEEK*
DNA starvation/stationary phase protection protein (DPS)		
151	dps(te)-rTT 1	*SATTTLKEQVLTTLKREQANAVVMYLNYKKYHWLTYGPLFRDLHLLFEEQGSEVFAMIDELA ERSLMLDGQPVADPADYLKVATVTPSSGQLTVKQMIEEAIANHELIITEMHQDAEIATEAGD IGTADLYTRLVQTHQKHRWFLKEFLAKGDGLVS*ggsggggsggKNLDCWVDNEEDIDVILKKS TILNLDINNDIISDISGFNS**A**VITYPDAQLVPGINGKAIHLVNNE**A**SEVIVHKAMDIEYNDM FN**Q**FTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSA GEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGVLMGSAEITGLGAIREDNQI TLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITFLRDFWGNPLRYDTEYYLIP VASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNEIDSFVKSGD FIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLK LYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND
152	dps(te)-rTT 2	*SATTTLKEQVLTTLKREQANAVVMYLNYKKYHWLTYGPLFRDLHLLFEEQGSEVFAMIDELA ERSLMLDGQPVADPADYLKVATVTPSSGQLTVKQMIEEAIANHELIITEMHQDAEIATEAGD IGTADLYTRLVQTHQKHRWFLKEFLAKGDGLVS*ggsggggsgggsggsgKNLDCWVDNEEDID VILKKSTILNLDINNDIISDISGFNS**A**VITYPDAQLVPGINGKAIHLVNNE**A**SEVIVHKAMD IEYNDMFN**Q**FTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIW TLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGVLMGSAEITGLGAI REDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITFLRDFWGNPLRYDT EYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNEIDS FVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDLKT YSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGW TND
153	dps(te)-rTT 3	*SATTTLKEQVLTTLKREQANAVVMYLNYKKYHWLTYGPLFRDLHLLFEEQGSEVFAMIDELA ERSLMLDGQPVADPADYLKVATVTPSSGQLTVKQMIEEAIANHELIITEMHQDAEIATEAGD IGTADLYTRLVQTHQKHRWFLKEFLAKGDGLVS*ggdsggdgggdggdgKNLDCWVDNEEDID VILKKSTILNLDINNDIISDISGFNS**A**VITYPDAQLVPGINGKAIHLVNNE**A**SEVIVHKAMD IEYNDMFN**Q**FTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIW TLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGVLMGSAEITGLGAI REDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITFLRDFWGNPLRYDT EYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNEIDS FVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDLKT YSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGW TND
154	rTT-dps(te) 1	NLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNS**A**VITYPDAQLVPGINGKAIHLVN NE**A**SEVIVHKAMDIEYNDMFN**Q**FTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGS GWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGV

		LMGSAEITGLGAIREDN**Q**ITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITF LRDFWGNPLRYDTEYYLIPVASSSKDVQLK**Q**ITDYMYLTNAPSYTNGKLNIYYRRLYNGLKF IIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLY KKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKIL GCDWYFVPTDEGWTNDggsgg*SATTTLKEQVLTTLKREQANAVVMYLNYKKYHWLTYGPLFR* *DLHLLFEEQGSEVFAMIDELAERSLMLDGQPVADPADYLKVATVTPSSGQLTVKQMIEEAIA* *NHELIITEMHQDAEIATEAGDIGTADLYTRLVQTHQKHRWFLKEFLAKGDGLVS*
155	rTT-dps(te) 2	NLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNS**A**VITYPDAQLVPGINGKAIHLVN NE**A**SEVIVHKAMDIEYNDMFN**Q**FTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGS GWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGV LMGSAEITGLGAIREDN**Q**ITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITF LRDFWGNPLRYDTEYYLIPVASSSKDVQLK**Q**ITDYMYLTNAPSYTNGKLNIYYRRLYNGLKF IIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLY KKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKIL GCDWYFVPTDEGWTNDggsggsgg*SATTTLKEQVLTTLKREQANAVVMYLNYKKYHWLTYGP* *LFRDLHLLFEEQGSEVFAMIDELAERSLMLDGQPVADPADYLKVATVTPSSGQLTVKQMIEE* *AIANHELIITEMHQDAEIATEAGDIGTADLYTRLVQTHQKHRWFLKEFLAKGDGLVS*
156	rTT-dps(te) 3	NLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNS**A**VITYPDAQLVPGINGKAIHLVN NE**A**SEVIVHKAMDIEYNDMFN**Q**FTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGS GWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGV LMGSAEITGLGAIREDN**Q**ITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITF LRDFWGNPLRYDTEYYLIPVASSSKDVQLK**Q**ITDYMYLTNAPSYTNGKLNIYYRRLYNGLKF IIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLY KKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKIL GCDWYFVPTDEGWTNDggdggdgg*SATTTLKEQVLTTLKREQANAVVMYLNYKKYHWLTYGP* *LFRDLHLLFEEQGSEVFAMIDELAERSLMLDGQPVADPADYLKVATVTPSSGQLTVKQMIEE* *AIANHELIITEMHQDAEIATEAGDIGTADLYTRLVQTHQKHRWFLKEFLAKGDGLVS*
157	dps(kr)-rTT 1	*TTIHDVQTTGLTQDAVTGFDASSRLNAGLQEVLVDLTALHLQGKQAHWNIVGENWRDLHLQL* *DTLVEAARGFSDDVAERMRAVGGVPDARPQTVAASRIGDVGPDEIDTRACVEAIVALVRHTV* *DTIRRVHDPIDAEDPASADLLHAITLELEKQAWMIGSENRSPRR*ggsggKNLDCWVDNEEDI DVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHLVNNEASEVIVHKAM DIEYNDMFNQFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLI WTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGVLMGSAEITGLGA IREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITFLRDFWGNPLRYD TEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNEID SFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDLK TYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEG WTND
158	dps(kr)-rTT 2	*TTIHDVQTTGLTQDAVTGFDASSRLNAGLQEVLVDLTALHLQGKQAHWNIVGENWRDLHLQL* *DTLVEAARGFSDDVAERMRAVGGVPDARPQTVAASRIGDVGPDEIDTRACVEAIVALVRHTV* *DTIRRVHDPIDAEDPASADLLHAITLELEKQAWMIGSENRSPRR*ggsgggsggKNLDCWVDN EEDIDVILKKSTILNLDINNDIISDISGFNS**A**VITYPDAQLVPGINGKAIHLVNNE**A**SEVIV HKAMDIEYNDMFN**Q**FTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKG NNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGVLMGSAEIT GLGAIREDN**Q**ITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITFLRDFWGNP LRYDTEYYLIPVASSSKDVQLK**Q**ITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPN NEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKL RDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFVP TDEGWTND
159	dps(kr)-rTT 3	*TTIHDVQTTGLTQDAVTGFDASSRLNAGLQEVLVDLTALHLQGKQAHWNIVGENWRDLHLQL* *DTLVEAARGFSDDVAERMRAVGGVPDARPQTVAASRIGDVGPDEIDTRACVEAIVALVRHTV* *DTIRRVHDPIDAEDPASADLLHAITLELEKQAWMIGSENRSPRR*ggsgggsgggsggsgKNL DCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNS**A**VITYPDAQLVPGINGKAIHLVNNE **A**SEVIVHKAMDIEYNDMFN**Q**FTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGW SVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGVLM GSAEITGLGAIREDN**Q**ITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITFLR DFWGNPLRYDTEYYLIPVASSSKDVQLK**Q**ITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFII KRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKK MEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGC DWYFVPTDEGWTND
160	rTT-dps(kr) 1	NLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNS**A**VITYPDAQLVPGINGKAIHLVN NE**A**SEVIVHKAMDIEYNDMFN**Q**FTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGS

		GWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGV LMGSAEITGLGAIREDN**Q**ITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITF LRDFWGNPLRYDTEYYLIPVASSSKDVQLK**Q**ITDYMYLTNAPSYTNGKLNIYYRRLYNGLKF IIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLY KKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKIL GCDWYFVPTDEGWTNDggsgg*TTIHDVQTTGLTQDAVTGFDASSRLNAGLQEVLVDLTALHL QGKQAHWNIVGENWRDLHLQLDTLVEAARGFSDDVAERMRAVGGVPDARPQTVAASRIGDVG PDEIDTRACVEAIVALVRHTVDTIRRVHDPIDAEDPASADLLHAITLELEKQAWMIGSENRS PRRR*
161	rTT-dps(kr) 2	NLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNS**A**VITYPDAQLVPGINGKAIHLVN NE**A**SEVIVHKAMDIEYNDMFN**Q**FTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGS GWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGV LMGSAEITGLGAIREDN**Q**ITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITF LRDFWGNPLRYDTEYYLIPVASSSKDVQLK**Q**ITDYMYLTNAPSYTNGKLNIYYRRLYNGLKF IIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLY KKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKIL GCDWYFVPTDEGWTNDggsgggsgg*TTIHDVQTTGLTQDAVTGFDASSRLNAGLQEVLVDLT ALHLQGKQAHWNIVGENWRDLHLQLDTLVEAARGFSDDVAERMRAVGGVPDARPQTVAASRI GDVGPDEIDTRACVEAIVALVRHTVDTIRRVHDPIDAEDPASADLLHAITLELEKQAWMIGS ENRSPRRR*
162	dps(np)-rTT 1	*SETQTLLRNFGNVYDNPVLLDRSVTAPVTEGFNVVLASFQALYLQYQKHHFVVEGSEFYSLH EFFNEAYNQVQDHIHEIGERLDGLGGVPVATFSKLAELTCFEQESEGVYSSRQMVENDLAAE QAIIGVIRRQAAQAESLGDRGTRYLYEKILLKTEERAYHLSHFLAKDSLTLGFVQAAQS*ggs ggKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNS**A**VITYPDAQLVPGINGKAIH LVNNE**A**SEVIVHKAMDIEYNDMFN**Q**FTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLS IGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYI NGVLMGSAEITGLGAIREDN**Q**ITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLS ITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLK**Q**ITDYMYLTNAPSYTNGKLNIYYRRLYNG LKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGI PLYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKD KILGCDWYFVPTDEGWTND
163	dps(np)-rTT 2	*SETQTLLRNFGNVYDNPVLLDRSVTAPVTEGFNVVLASFQALYLQYQKHHFVVEGSEFYSLH EFFNEAYNQVQDHIHEIGERLDGLGGVPVATFSKLAELTCFEQESEGVYSSRQMVENDLAAE QAIIGVIRRQAAQAESLGDRGTRYLYEKILLKTEERAYHLSHFLAKDSLTLGFVQAAQS*ggs gggsggKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNS**A**VITYPDAQLVPGING KAIHLVNNE**A**SEVIVHKAMDIEYNDMFN**Q**FTVSFWLRVPKVSASHLEQYGTNEYSIISSMKK HSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSA NLYINGVLMGSAEITGLGAIREDN**Q**ITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYT SYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLK**Q**ITDYMYLTNAPSYTNGKLNIYYRR LYNGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYN APGIPLYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFN HLKDKILGCDWYFVPTDEGWTND
164	dps(np)-rTT 3	*SETQTLLRNFGNVYDNPVLLDRSVTAPVTEGFNVVLASFQALYLQYQKHHFVVEGSEFYSLH EFFNEAYNQVQDHIHEIGERLDGLGGVPVATFSKLAELTCFEQESEGVYSSRQMVENDLAAE QAIIGVIRRQAAQAESLGDRGTRYLYEKILLKTEERAYHLSHFLAKDSLTLGFVQAAQS*ggs gggsgggsggsgKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNS**A**VITYPDAQL VPGINGKAIHLVNNE**A**SEVIVHKAMDIEYNDMFN**Q**FTVSFWLRVPKVSASHLEQYGTNEYSI ISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITN DRLSSANLYINGVLMGSAEITGLGAIREDN**Q**ITLKLDRCNNNNQYVSIDKFRIFCKALNPKE IEKLYTSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLK**Q**ITDYMYLTNAPSYTNGKL NIYYRRLYNGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRI LRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIA SNWYFNHLKDKILGCDWYFVPTDEGWTND
165	rTT-dps(np) 1	NLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNS**A**VITYPDAQLVPGINGKAIHLVN NE**A**SEVIVHKAMDIEYNDMFN**Q**FTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGS GWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGV LMGSAEITGLGAIREDN**Q**ITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITF LRDFWGNPLRYDTEYYLIPVASSSKDVQLK**Q**ITDYMYLTNAPSYTNGKLNIYYRRLYNGLKF IIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLY KKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKIL GCDWYFVPTDEGWTNDggsgg*MSETQTLLRNFGNVYDNPVLLDRSVTAPVTEGFNVVLASFQ ALYLQYQKHHFVVEGSEFYSLHEFFNEAYNQVQDHIHEIGERLDGLGGVPVATFSKLAELTC*

		FEQESEGVYSSRQMVENDLAAEQAIIGVIRRQAAQAESLGDRGTRYLYEKILLKTEERAYHL SHFLAKDSLTLGFVQAAQS
166	rTT-dps(np) 2	NLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNS**A**VITYPDAQLVPGINGKAIHLVN NE**A**SEVIVHKAMDIEYNDMFN**Q**FTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGS GWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGV LMGSAEITGLGAIREDN**Q**ITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITF LRDFWGNPLRYDTEYYLIPVASSSKDVQLK**Q**ITDYMYLTNAPSYTNGKLNIYYRRLYNGLKF IIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLY KKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKIL GCDWYFVPTDEGWTNDggsgggsgg*MSETQTLLRNFGNVYDNPVLLDRSVTAPVTEGFNVVL ASFQALYLQYQKHHFVVEGSEFYSLHEFFNEAYNQVQDHIHEIGERLDGLGGVPVATFSKLA ELTCFEQESEGVYSSRQMVENDLAAEQAIIGVIRRQAAQAESLGDRGTRYLYEKILLKTEER AYHLSHFLAKDSLTLGFVQAAQS*
Bacteriophage Q Beta Capsid, Chain A		
167	rTT-qbeta 1	NLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNS**A**VITYPDAQLVPGINGKAIHLVN NE**A**SEVIVHKAMDIEYNDMFN**Q**FTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGS GWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGV LMGSAEITGLGAIREDN**Q**ITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITF LRDFWGNPLRYDTEYYLIPVASSSKDVQLK**Q**ITDYMYLTNAPSYTNGKLNIYYRRLYNGLKF IIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLY KKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKIL GCDWYFVPTDEGWTNDggsggsgsgg*AKLETVTLGNIGKDGKQTLVLNPRGVNPTNGVASLS QAGAVPALEKRVTVSVSQPSRNRKNYKVQVKIQNPTACTANG**A**CDPSVTRQAYADVTFSFTQ YSTDEERAFVRTELAALLASPLLIDAIDQLNPAY*
168	rTT-qbeta 2	NLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNS**A**VITYPDAQLVPGINGKAIHLVN NE**A**SEVIVHKAMDIEYNDMFN**Q**FTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGS GWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGV LMGSAEITGLGAIREDN**Q**ITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITF LRDFWGNPLRYDTEYYLIPVASSSKDVQLK**Q**ITDYMYLTNAPSYTNGKLNIYYRRLYNGLKF IIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLY KKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKIL GCDWYFVPTDEGWTNDggdgssgdgg*AKLETVTLGNIGKDGKQTLVLNPRGVNPTNGVASLS QAGAVPALEKRVTVSVSQPSRNRKNYKVQVKIQNPTACTANG**A**CDPSVTRQAYADVTFSFTQ YSTDEERAFVRTELAALLASPLLIDAIDQLNPAY*
169	rTT-qbeta 3	NLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNS**A**VITYPDAQLVPGINGKAIHLVN NE**A**SEVIVHKAMDIEYNDMFN**Q**FTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGS GWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGV LMGSAEITGLGAIREDN**Q**ITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITF LRDFWGNPLRYDTEYYLIPVASSSKDVQLK**Q**ITDYMYLTNAPSYTNGKLNIYYRRLYNGLKF IIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLY KKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKIL GCDWYFVPTDEGWTNDggsgggsgggsggsg*AKLETVTLGNIGKDGKQTLVLNPRGVNPTNG VASLSQAGAVPALEKRVTVSVSQPSRNRKNYKVQVKIQNPTACTANG**A**CDPSVTRQAYADVT FSFTQYSTDEERAFVRTELAALLASPLLIDAIDQLNPAY*
170	rTT-qbeta 4	NLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNS**A**VITYPDAQLVPGINGKAIHLVN NE**A**SEVIVHKAMDIEYNDMFN**Q**FTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGS GWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGV LMGSAEITGLGAIREDN**Q**ITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITF LRDFWGNPLRYDTEYYLIPVASSSKDVQLK**Q**ITDYMYLTNAPSYTNGKLNIYYRRLYNGLKF IIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLY KKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKIL GCDWYFVPTDEGWTNDggdsggdgggdggdg*AKLETVTLGNIGKDGKQTLVLNPRGVNPTNG VASLSQAGAVPALEKRVTVSVSQPSRNRKNYKVQVKIQNPTACTANG**A**CDPSVTRQAYADVT FSFTQYSTDEERAFVRTELAALLASPLLIDAIDQLNPAY*
171	rTT-qbeta 5	NLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNS**A**VITYPDAQLVPGINGKAIHLVN NE**A**SEVIVHKAMDIEYNDMFN**Q**FTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGS GWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGV LMGSAEITGLGAIREDN**Q**ITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITF LRDFWGNPLRYDTEYYLIPVASSSKDVQLK**Q**ITDYMYLTNAPSYTNGKLNIYYRRLYNGLKF IIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLY KKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKIL

		GCDWYFVPTDEGWTNDggsggpppppgsggsg*AKLETVTLGNIGKDGKQTLVLNPRGVNPTNG VASLSQAGAVPALEKRVTVSVSQPSRNRKNYKVQVKIQNPTACTANG***A***CDPSVTRQAYADVT FSFTQYSTDEERAFVRTELAALLASPLLIDAIDQLNPAY*
CaMKIIa (12-mer) C-term fragment (5U6Y)		
172	CRM-5U6Y 1	GADDVVDSSKSFVMENFSSYHGTKPGYVDSIQKGIQKPKSGTQGNYDDDWKEFYSTDNKYDA AGYSVDNENPLSGKAGGVVKVTYPGLTKVLALKVDNAETIKKELGLSLTEPLMEQVGTEEFI KRFGDGASRVVLSLPFAEGSSSVEYINNWEQAKALSVELEINFETRGKRGQDAMYEYMAQAC AGNRVRRSVGSSLSCINLDWDVIRDKTKTKIESLKEHGPIKNKMSESPNKTVSEEKAKQYLE EFHQTALEHPELSELKTVTGTNPVFAGANYAAWAVNVAQVIDSETADNLEKTTAALSILPGI GSVMGIADGAVHHNTEEIVAQSIALSSLMVAQAIPLVGELVDIGFAAYNFVESIINLFQVVH NSYNRPAYSPGHKTQPFLHDGYAVSWNTVEDSIIRTGFQGESGHDIKITAENTPLPIAGVLL PTIPGKLDVNKSKTHISVNGRKIRMRCRAIDGDVTFCRPKSPVYVGNGVHANLHVAFHRSSS EKIHSNEISSDSIGVLGYQKTVDHTKVNSKLSLFFEIKS*ggksggnkksdgvkessestnta iededtkvrkqeiikvteqlieaisngdfesytkmcdpgmtafepealgnlvegldfhrfyf enlwsrnskpvhttilnphihlmgdesaciayiritqyldaggiprtaqseetrvwhrrdgk wqivhfhrsga*
173	CRM-5U6Y 2	GADDVVDSSKSFVMENFSSYHGTKPGYVDSIQKGIQKPKSGTQGNYDDDWKEFYSTDNKYDA AGYSVDNENPLSGKAGGVVKVTYPGLTKVLALKVDNAETIKKELGLSLTEPLMEQVGTEEFI KRFGDGASRVVLSLPFAEGSSSVEYINNWEQAKALSVELEINFETRGKRGQDAMYEYMAQAC AGNRVRRSVGSSLSCINLDWDVIRDKTKTKIESLKEHGPIKNKMSESPNKTVSEEKAKQYLE EFHQTALEHPELSELKTVTGTNPVFAGANYAAWAVNVAQVIDSETADNLEKTTAALSILPGI GSVMGIADGAVHHNTEEIVAQSIALSSLMVAQAIPLVGELVDIGFAAYNFVESIINLFQVVH NSYNRPAYSPGHKTQPFLHDGYAVSWNTVEDSIIRTGFQGESGHDIKITAENTPLPIAGVLL PTIPGKLDVNKSKTHISVNGRKIRMRCRAIDGDVTFCRPKSPVYVGNGVHANLHVAFHRSSS EKIHSNEISSDSIGVLGYQKTVDHTKVNSKLSLFFEIKS*gvkessestntaiededtkvrkq eiikvteqlieaisngdfesytkmcdpgmtafepealgnlvegldfhrfyfenlwsrnskpv httilnphihlmgdesaciayiritqyldaggiprtaqseetrvwhrrdgkwqivhfhrsga*
174	CRM-5U6Y 3	GADDVVDSSKSFVMENFSSYHGTKPGYVDSIQKGIQKPKSGTQGNYDDDWKEFYSTDNKYDA AGYSVDNENPLSGKAGGVVKVTYPGLTKVLALKVDNAETIKKELGLSLTEPLMEQVGTEEFI KRFGDGASRVVLSLPFAEGSSSVEYINNWEQAKALSVELEINFETRGKRGQDAMYEYMAQAC AGNRVRRSVGSSLSCINLDWDVIRDKTKTKIESLKEHGPIKNKMSESPNKTVSEEKAKQYLE EFHQTALEHPELSELKTVTGTNPVFAGANYAAWAVNVAQVIDSETADNLEKTTAALSILPGI GSVMGIADGAVHHNTEEIVAQSIALSSLMVAQAIPLVGELVDIGFAAYNFVESIINLFQVVH NSYNRPAYSPGHKTQPFLHDGYAVSWNTVEDSIIRTGFQGESGHDIKITAENTPLPIAGVLL PTIPGKLDVNKSKTHISVNGRKIRMRCRAIDGDVTFCRPKSPVYVGNGVHANLHVAFHRSSS EKIHSNEISSDSIGVLGYQKTVDHTKVNSKLSLFFEIKS*gstntaiededtkvrkqeiikvt eqlieaisngdfesytkmcdpgmtafepealgnlvegldfhrfyfenlwsrnskpvhttiln phihlmgdesaciayiritqyldaggiprtaqseetrvwhrrdgkwqivhfhrsga*
175	HID-5U6Y 1	SNMANTQMKSDKIIIAHRGASGYLPEHTLESKALAFAQQADYLEQDLAMTKDGRLVVIHDHF LDGLTDVAKKFPHRHRKDGRYYVIDFTLKEIQSLEMTENFETKDGKQAQVYPNRFPLWKSHF RIHTFEDEIEFIQGLEKSTGKKVGIYPEIKAPWFHHQNGKDIAAETLKVLKKYGYDKKTDMV YLQTFDFNELKRIKTELLPQMGMDLKLVQLIAYTDWKETQEKDPKGYWVNYNYDWMFKPGAM AEVVKYADGVGPGWYMLVNKEESKPDNIVYTPLVKELAQYNVEVHPYTVRKDALPEFFTDVN QMYDALLNKSGATGVFTDFPDTGVEFLKGIK*ggksggnkksdgvkessestntaiededtkv rkqeiikvteqlieaisngdfesytkmcdpgmtafepealgnlvegldfhrfyfenlwsrns kpvhttilnphihlmgdesaciayiritqyldaggiprtaqseetrvwhrrdgkwqivhfhr sga*
176	HID-5U6Y 2	SNMANTQMKSDKIIIAHRGASGYLPEHTLESKALAFAQQADYLEQDLAMTKDGRLVVIHDHF LDGLTDVAKKFPHRHRKDGRYYVIDFTLKEIQSLEMTENFETKDGKQAQVYPNRFPLWKSHF RIHTFEDEIEFIQGLEKSTGKKVGIYPEIKAPWFHHQNGKDIAAETLKVLKKYGYDKKTDMV YLQTFDFNELKRIKTELLPQMGMDLKLVQLIAYTDWKETQEKDPKGYWVNYNYDWMFKPGAM AEVVKYADGVGPGWYMLVNKEESKPDNIVYTPLVKELAQYNVEVHPYTVRKDALPEFFTDVN QMYDALLNKSGATGVFTDFPDTGVEFLKGIK*gvkessestntaiededtkvrkqeiikvteq lieaisngdfesytkmcdpgmtafepealgnlvegldfhrfyfenlwsrnskpvhttilnph ihlmgdesaciayiritqyldaggiprtaqseetrvwhrrdgkwqivhfhrsga*
177	HID-5U6Y 3	SNMANTQMKSDKIIIAHRGASGYLPEHTLESKALAFAQQADYLEQDLAMTKDGRLVVIHDHF LDGLTDVAKKFPHRHRKDGRYYVIDFTLKEIQSLEMTENFETKDGKQAQVYPNRFPLWKSHF RIHTFEDEIEFIQGLEKSTGKKVGIYPEIKAPWFHHQNGKDIAAETLKVLKKYGYDKKTDMV YLQTFDFNELKRIKTELLPQMGMDLKLVQLIAYTDWKETQEKDPKGYWVNYNYDWMFKPGAM AEVVKYADGVGPGWYMLVNKEESKPDNIVYTPLVKELAQYNVEVHPYTVRKDALPEFFTDVN QMYDALLNKSGATGVFTDFPDTGVEFLKGIK*gstntaiededtkvrkqeiikvteqlieais*

		ngdfesytkmcdpgmtafepealgnlvegldfhrfyfenlwsrnskpvhttilnphihlmgd esaciayiritqyldaggiprtaqseetrvwhrrdgkwqivhfhrsga
Phosphopantetheine Adenylyltransferase (6ccq)		
178	CRM-6CCQ-rTT	GADDVVDSSKSFVMENFSSYHGTKPGYVDSIQKGIQKPKSGTQGNYDDDWKEFYSTDNKYDA AGYSVDNENPLSGKAGGVVKVTYPGLTKVLALKVDNAETIKKELGLSLTEPLMEQVGTEEFI KRFGDGASRVVLSLPFAEGSSSVEYINNWEQAKALSVELEINFETRGKRGQDAMYEYMAQAC AGNRVRRSVGSSLSCINLDWDVIRDKTKTKIESLKEHGPIKNKMSESPNKTVSEEKAKQYLE EFHQTALEHPELSELKTVTGTNPVFAGANYAAWAVNVAQVIDSETADNLEKTTAALSILPGI GSVMGIADGAVHHNTEEIVAQSIALSSLMVAQAIPLVGELVDIGFAAYNFVESIINLFQVVH NSYNRPAYSPGHKTQPFLHDGYAVSWNTVEDSIIRTGFQGESGHDIKITAENTPLPIAGVLL PTIPGKLDVNKSKTHISVNGRKIRMRCRAIDGDVTFCRPKSPVYVGNGVHANLHVAFHRSSS EKIHSNEISSDSIGVLGYQKTVDHTKVNSKLSLFFEIKS*ggggsggggs MQKRAIYPGTFDP ITNGHIDIVTRATQMFDHVILAIAASPSKKPMFTLEERVALAQQATAHLGNVEVVGFSDLMA NFARNQHATVLIRGLRAVADFEYEMQLAHMNRHLMPELESVFLMPSKEWSFISSSLVKEVAR HQGDVTHFLPENVHQALMAKLAVD*ggggsggggs*MKNLDCWVDNEEDIDVILKKSTILNLDI NNDIISDISGFNSSVITYPDAQLVPGINGKAIHLVNNESSEVIVHKAMDIEYNDMFNNFTVS FWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQIT FRDLPDKFNAYLANKWVFITITNDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRC NNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKD VQLKNITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNEIDSFVKSGDFIKLYVS YNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKNA SLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND
179	HID-6CCQ-rTT	SNMANTQMKSDKIIIAHRGASGYLPEHTLESKALAFAQQADYLEQDLAMTKDGRLVVIHDHF LDGLTDVAKKFPHRHRKDGRYYVIDFTLKEIQSLEMTENFETKDGKQAQVYPNRFPLWKSHF RIHTFEDEIEFIQGLEKSTGKKVGIYPEIKAPWFHHQNGKDIAAETLKVLKKYGYDKKTDMV YLQTFDFNELKRIKTELLPQMGMDLKLVQLIAYTDWKETQEKDPKGYWVNYNYDWMFKPGAM AEVVKYADGVGPGWYMLVNKEESKPDNIVYTPLVKELAQYNVEVHPYTVRKDALPEFFTDVN QMYDALLNKSGATGVFTDFPDTGVEFLKGIK*ggggsggggs MQKRAIYPGTFDPITNGHIDI VTRATQMFDHVILAIAASPSKKPMFTLEERVALAQQATAHLGNVEVVGFSDLMANFARNQHA TVLIRGLRAVADFEYEMQLAHMNRHLMPELESVFLMPSKEWSFISSSLVKEVARHQGDVTHF LPENVHQALMAKLAVD*ggggsggggs*MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDI SGFNSSVITYPDAQLVPGINGKAIHLVNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKV SASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKF NAYLANKWVFITITNDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVS IDKFRIFCKALNPKEIEKLYTSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITD YMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIV GYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTH NGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND
T4 fibritin Foldon domain (Fd)		
180	Fd-rTT_degly	*GYIPEAPRDGQAYVRKDGEWVLLSTFL*GSGGGGGQMKNLDCWVDNEEDIDVILKKSTILNLD INNDIISDISGFNSAVITYPDAQLVPGINGKAIHLVNNEASEVIVHKAMDIEYNDMFNNFTV SFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQI TFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGVLMGSAEITGLGAIREDNQITLKLDR CNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITFLRDFWGNPLRYDTEYYLIPVASSSK DVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNEIDSFVKSGDFIKLYV SYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKQ ASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND
181	rTT_degly-Fd-TT_degly	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDgg*GS GYIPEAPRDGQAYVRKDGEWVLLSTFL*GSGGGGGQMKNLD CWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHLVNNEA SEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWS VSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGVLMG SAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITFLRD FWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIK

		RYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKM EAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCD WYFVPTDEGWTND
182	Fd- rTT_degly -K20	*GYIPEAPRDGQAYVRKDGEWVLLSTFL*GSGGGGGQMKNLDCWVDkEEDIDVILKKSTILNLD INkDIISDISkFNSAVITYPDAQLVPGINGKAIHLVNNEkSEVIVHKAMkIEYNDMFNNFTV SFWLRVPKVSASHLEQYGTNEYSIISSMKKHkkSIGSGWSVSLKGNNLIWTLKDSkGEVRQI TFRDLPkKFNAYLANKWVFITITNDRkSSANLYINGVLMGSAEITkLGAIREDNQITLKLDR CkNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITFLRDFWGNPLRYDTEYYLIPVASSSK DVQLKQITDYMYLTkAPSYTNGKLNIYYRRLYNGLKFIIKRYkPNNkIDSFVKSGDFIKLYV SYkNNEHIVGYPKDGNAFNkLDRILRVGYkAPkIPLYKKMEAVKLRDLKTYSVQLKLYDDKQ ASLGLVGTHNGQIGkDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND
183	rTT_degly -Fd- TT_degly- K20	MKNLDCWVDkEEDIDVILKKSTILNLDINkDIISDISkFNSAVITYPDAQLVPGINGKAIHL VNNEkSEVIVHKAMkIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHkkSI GSGWSVSLKGNNLIWTLKDSkGEVRQITFRDLPkKFNAYLANKWVFITITNDRkSSANLYIN GVLMGSAEITkLGAIREDNQITLKLDRCkNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTkAPSYTNGKLNIYYRRLYNGL KFIIKRYkPNNkIDSFVKSGDFIKLYVSYkNNEHIVGYPKDGNAFNkLDRILRVGYkAPkIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGkDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDggGS*GYIPEAPRDGQAYVRKDGEWVLLSTFL*GSGGGGGQMKNLD CWVDkEEDIDVILKKSTILNLDINkDIISDISkFNSAVITYPDAQLVPGINGKAIHLVNNEk SEVIVHKAMkIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHkkSIGSGWS VSLKGNNLIWTLKDSkGEVRQITFRDLPkKFNAYLANKWVFITITNDRkSSANLYINGVLMG SAEITkLGAIREDNQITLKLDRCkNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITFLRD FWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTkAPSYTNGKLNIYYRRLYNGLKFIIK RYkPNNkIDSFVKSGDFIKLYVSYkNNEHIVGYPKDGNAFNkLDRILRVGYkAPkIPLYKKM EAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGkDPNRDILIASNWYFNHLKDKILGCD WYFVPTDEGWTND
Hexamer		
184	rTT_degly - r8_linker -Fc- Hexamer	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDggaeaaakeaaakeaaakeaaakaleaeaaakeaaakeaaakea aakaEPKSCDKTHTCPKCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEV KFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVCLQDWLNGKEYKCKVSNKALPAPIEKT ISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPV LDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGKggsgg*PTLYNVS* *LVMSDTAGTCY*
185	rTT_degly - 12pa_link er-Fc- Hexamer	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDgspapapapapapapapapapapapapaEPKSCDKTHTCPKCPAPE LLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQ YNSTYRVVSVLTVCLQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDE LTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQ GNVFSCSVMHEALHNHYTQKSLSLSPGKggsgg*PTLYNVSLVMSDTAGTCY*
186	rTT_degly - r3_linker -Fc- Hexamer	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDggaeaaakeaaakeaaakaEPKSCDKTHTCPKCPAPELLGGPSV

		FLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRV VSVLTVCLQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVS LTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCS VMHEALHNHYTQKSLSLSPGKggsgg*PTLYNVSLVMSDTAGTCY*
187	rTT_degly-5ga_linker-Fc-Hexamer	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDgsgsgsgsgsasgasgEPKSCDKTHTCPKCPAPELLGGPSVFLF PPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSV LTVCLQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTC LVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMH EALHNHYTQKSLSLSPGKggsgg*PTLYNVSLVMSDTAGTCY*
188	rTT_degly-3f_linker-Fc-Hexamer	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDggggsggggsggggsEPKSCDKTHTCPKCPAPELLGGPSVFLFP PKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVL TVCLQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCL VKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHE ALHNHYTQKSLSLSPGKggsgg*PTLYNVSLVMSDTAGTCY*
189	rTT_degly-2rf_linker-Fc-Hexamer	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDgsaeaaakeaaakaEPKSCDKTHTCPKCPAPELLGGPSVFLFPP KPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLT VCLQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLV KGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEA LHNHYTQKSLSLSPGKggsgg*PTLYNVSLVMSDTAGTCY*
190	rTT_degly-1f_linker-Fc-Hexamer	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDggggsasgEPKSCDKTHTCPKCPAPELLGGPSVFLFPPKPKDTL MISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVCLQDW LNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPS DIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYT QKSLSLSPGKggsgg*PTLYNVSLVMSDTAGTCY*
191	rTT_degly-1rf_liner-Fc-Hexamer	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDpapapasgEPKSCDKTHTCPKCPAPELLGGPSVFLFPPKPKDTL MISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVCLQDW LNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPS

		DIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYT QKSLSLSPGKggsgg*PTLYNVSLVMSDTAGTCY*
DIHYDROLIPOYL TRANSACETYLASE (e2p)		
192	CRM-e2p 1	GADDVVDSSKSFVMENFSSYHGTKPGYVDSIQKGIQKPKSGTQGNYDDDWKEFYSTDNKYDA AGYSVDNENPLSGKAGGVVKVTYPGLTKVLALKVDNAETIKKELGLSLTEPLMEQVGTEEFI KRFGDGASRVVLSLPFAEGSSSVEYINNWEQAKALSVELEINFETRGKRGQDAMYEYMAQAC AGNRVRRSVGSSLSCINLDWDVIRDKTKTKIESLKEHGPIKNKMSESPNKTVSEEKAKQYLE EFHQTALEHPELSELKTVTGTNPVFAGANYAAWAVNVAQVIDSETADNLEKTTAALSILPGI GSVMGIADGAVHHNTEEIVAQSIALSSLMVAQAIPLVGELVDIGFAAYNFVESIINLFQVVH NSYNRPAYSPGHKTQPFLHDGYAVSWNTVEDSIIRTGFQGESGHDIKITAENTPLPIAGVLL PTIPGKLDVNKSKTHISVNGRKIRMRCRAIDGDVTFCRPKSPVYVGNGVHANLHVAFHRSSS EKIHSNEISSDSIGVLGYQKTVDHTKVNSKLSLFFEIK*ggggsggggsGAAAKPATTEGEFP ETREKMSGIRRAIAKAMVHSKHTAPHVTLMDEADVTKLVAHRKKFKAIAAEKGIKLTFLPYV VKALVSALREYPVLNT**A**IDDETEEIIQKHYYNIGIAADTDRGLLVPVIKHADRKPIFALAQE INELAEKARDGKLTPGEMKGASCTITNIGSAGGQWFTPVINHPEVAILGIGRIAEKPIVRDG EIVAAPMLALSLSFDHRMIDGATAQKALNHIKRLLSDPELLLM***
193	HID-e2p	SNMANTQMKSDKIIIAHRGASGYLPEHTLESKALAFAQQADYLEQDLAMTKDGRLVVIHDHF LDGLTDVAKKFPHRHRKDGRYYVIDFTLKEIQSLEMTENFETKDGKQAQVYPNRFPLWKSHF RIHTFEDEIEFIQGLEKSTGKKVGIYPEIKAPWFHHQNGKDIAAETLKVLKKYGYDKKTDMV YLQTFDFNELKRIKTELLPQMGMDLKLVQLIAYTDWKETQEKDPKGYWVNYNYDWMFKPGAM AEVVKYADGVGPGWYMLVNKEESKPDNIVYTPLVKELAQYNVEVHPYTVRKDALPEFFTDVN QMYDALLNKSGATGVFTDFPDTGVEFLKGIK*ggggsggggsGAAAKPATTEGEFPETREKMS GIRRAIAKAMVHSKHTAPHVTLMDEADVTKLVAHRKKFKAIAAEKGIKLTFLPYVVKALVSA LREYPVLNT**A**IDDETEEIIQKHYYNIGIAADTDRGLLVPVIKHADRKPIFALAQEINELAEK ARDGKLTPGEMKGASCTITNIGSAGGQWFTPVINHPEVAILGIGRIAEKPIVRDGEIVAAPM LALSLSFDHRMIDGATAQKALNHIKRLLSDPELLLM***
194	CRM-e2p 2	GADDVVDSSKSFVMENFSSYHGTKPGYVDSIQKGIQKPKSGTQGNYDDDWKEFYSTDNKYDA AGYSVDNENPLSGKAGGVVKVTYPGLTKVLALKVDNAETIKKELGLSLTEPLMEQVGTEEFI KRFGDGASRVVLSLPFAEGSSSVEYINNWEQAKALSVELEINFETRGKRGQDAMYEYMAQAC AGNRVRRSVGSSLSCINLDWDVIRDKTKTKIESLKEHGPIKNKMSESPNKTVSEEKAKQYLE EFHQTALEHPELSELKTVTGTNPVFAGANYAAWAVNVAQVIDSETADNLEKTTAALSILPGI GSVMGIADGAVHHNTEEIVAQSIALSSLMVAQAIPLVGELVDIGFAAYNFVESIINLFQVVH NSYNRPAYSPGHKTQPFLHDGYAVSWNTVEDSIIRTGFQGESGHDIKITAENTPLPIAGVLL PTIPGKLDVNKSKTHISVNGRKIRMRCRAIDGDVTFCRPKSPVYVGNGVHANLHVAFHRSSS EKIHSNEISSDSIGVLGYQKTVDHTKVNSKLSLFFEIKS*ggggsggggsGAAAKPATTEGEF PETREKMSGIRRAIAKAMVHSKHTAPHVTLMDEADVTKLVAHRKKFKAIAAEKGIKLTFLPY VVKALVSALREYPVLNT**A**IDDETEEIIQKHYYNIGIAADTDRGLLVPVIKHADRKPIFALAQ EINELAEKARDGKLTPGEMKGASCTITNIGSAGGQWFTPVINHPEVAILGIGRIAEKPIVRD GEIVAAPMLALSLSFDHRMIDGATAQKALNHIKRLLSDPELLLM***
195	rTT_degly -10f-E2p	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTND*ggggsggggsggggsggggsaeaaakeaaakaggggsggggsgg ggsggggs**AAAKPATTEGEFPETREKMSGIRRAIAKAMVHSKHTAPHVTLMDEADVTKLVAH RKKFKAIAAEKGIKLTFLPYVVKALVSALREYPVLNT**A**IDDETEEIIQKHYYNIGIAADTDR GLLVPVIKHADRKPIFALAQEINELAEKARDGKLTPGEMKGASCTITNIGSAGGQWFTPVIN HPEVAILGIGRIAEKPIVRDGEIVAAPMLALSLSFDHRMIDGATAQKALNHIKRLLSDPELL LM***
196	rTT_degly -8f-E2p	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTND*ggggsggggsggggsggggsggggsggggsggggsggggs**AAAK PATTEGEFPETREKMSGIRRAIAKAMVHSKHTAPHVTLMDEADVTKLVAHRKKFKAIAAEKG***

		*IKLTFLPYVVKALVSALREYPVLNT**A**IDDETEEIIQKHYYNIGIAADTDRGLLVPVIKHADR KPIFALAQEINELAEKARDGKLTPGEMKGASCTITNIGSAGGQWFTPVINHPEVAILGIGRI AEKPIVRDGEIVAAPMLALSLSFDHRMIDGATAQKALNHIKRLLSDPELLLM*
197	rTT_degly -4f-E2p	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDggggsggggsggggsggggs*AAAKPATTEGEFPETREKMSGIRR AIAKAMVHSKHTAPHVTLMDEADVTKLVAHRKKFKAIAAEKGIKLTFLPYVVKALVSALREY PVLNT**A**IDDETEEIIQKHYYNIGIAADTDRGLLVPVIKHADRKPIFALAQEINELAEKARDG KLTPGEMKGASCTITNIGSAGGQWFTPVINHPEVAILGIGRIAEKPIVRDGEIVAAPMLALS LSFDHRMIDGATAQKALNHIKRLLSDPELLLM*
198	rTT_degly -3f-E2p	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDggggsggggsggggs*AAAKPATTEGEFPETREKMSGIRRAIAKA MVHSKHTAPHVTLMDEADVTKLVAHRKKFKAIAAEKGIKLTFLPYVVKALVSALREYPVLNT **A**IDDETEEIIQKHYYNIGIAADTDRGLLVPVIKHADRKPIFALAQEINELAEKARDGKLTPG EMKGASCTITNIGSAGGQWFTPVINHPEVAILGIGRIAEKPIVRDGEIVAAPMLALSLSFDH RMIDGATAQKALNHIKRLLSDPELLLM*
199	rTT_degly -2f-E2p	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDggggsggggs*AAAKPATTEGEFPETREKMSGIRRAIAKAMVHSK HTAPHVTLMDEADVTKLVAHRKKFKAIAAEKGIKLTFLPYVVKALVSALREYPVLNT**A**IDDE TEEIIQKHYYNIGIAADTDRGLLVPVIKHADRKPIFALAQEINELAEKARDGKLTPGEMKGA SCTITNIGSAGGQWFTPVINHPEVAILGIGRIAEKPIVRDGEIVAAPMLALSLSFDHRMIDG ATAQKALNHIKRLLSDPELLLM*
200	rTT_degly -1f-E2p	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDggggsasg*AAAKPATTEGEFPETREKMSGIRRAIAKAMVHSKHT APHVTLMDEADVTKLVAHRKKFKAIAAEKGIKLTFLPYVVKALVSALREYPVLNT**A**IDDETE EIIQKHYYNIGIAADTDRGLLVPVIKHADRKPIFALAQEINELAEKARDGKLTPGEMKGASC TITNIGSAGGQWFTPVINHPEVAILGIGRIAEKPIVRDGEIVAAPMLALSLSFDHRMIDGAT AQKALNHIKRLLSDPELLLM*
201	rTT_degly -r8-E2p	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDggaeaaakeaaakeaaakeaaakaleaeaaakeaaakeaaakea aakasg*AAAKPATTEGEFPETREKMSGIRRAIAKAMVHSKHTAPHVTLMDEADVTKLVAHRK KFKAIAAEKGIKLTFLPYVVKALVSALREYPVLNT**A**IDDETEEIIQKHYYNIGIAADTDRGL*

		LVPVIKHADRKPIFALAQEINELAEKARDGKLTPGEMKGASCTITNIGSAGGQWFTPVINHP EVAILGIGRIAEKPIVRDGEIVAAPMLALSLSFDHRMIDGATAQKALNHIKRLLSDPELLLM
202	rTT_degly -r12-E2p	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDgsapapapapapapapapapapapapasg*AAAKPATTEGEFPET REKMSGIRRAIAKAMVHSKHTAPHVTLMDEADVTKLVAHRKKFKAIAAEKGIKLTFLPYVVK ALVSALREYPVLNT***A***IDDETEEIIQKHYYNIGIAADTDRGLLVPVIKHADRKPIFALAQEIN ELAEKARDGKLTPGEMKGASCTITNIGSAGGQWFTPVINHPEVAILGIGRIAEKPIVRDGEI VAAPMLALSLSFDHRMIDGATAQKALNHIKRLLSDPELLLM*
203	rTT_degly -r3-E2p	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDgaeaaakeaaakeaaakasg*AAAKPATTEGEFPETREKMSGIRR AIAKAMVHSKHTAPHVTLMDEADVTKLVAHRKKFKAIAAEKGIKLTFLPYVVKALVSALREY PVLNT***A***IDDETEEIIQKHYYNIGIAADTDRGLLVPVIKHADRKPIFALAQEINELAEKARDG KLTPGEMKGASCTITNIGSAGGQWFTPVINHPEVAILGIGRIAEKPIVRDGEIVAAPMLALS LSFDHRMIDGATAQKALNHIKRLLSDPELLLM*
204	rTT_degly -2rf-E2p	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSAVITYPDAQLVPGINGKAIHL VNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDgsaeaaakeaaakasg*AAAKPATTEGEFPETREKMSGIRRAIAK AMVHSKHTAPHVTLMDEADVTKLVAHRKKFKAIAAEKGIKLTFLPYVVKALVSALREYPVLN T***A***IDDETEEIIQKHYYNIGIAADTDRGLLVPVIKHADRKPIFALAQEINELAEKARDGKLTP GEMKGASCTITNIGSAGGQWFTPVINHPEVAILGIGRIAEKPIVRDGEIVAAPMLALSLSFD HRMIDGATAQKALNHIKRLLSDPELLLM*
Glutamate Synthase, Chain A (1f52)		
205	6H-3C-HiD-1f52	Snmantqmksdkiiiahrgasgylpehtleskalafaqqadyleqdlamtkdgrlvvihdhf ldgltdvakkfphrhrkdgryyvidftlkeiqslemtenfetkdgkqaqvypnrfplwkshf rihtfedeiefiqglekstgkkvgiypeikapwfhhqngkdiaaetlkvlkkygydkktdmv ylqtfdfnelkriktellpqmgmdlklvqliaytdwketqekdpkgywvnynydwmfkpgam aevvkyadgvgpgwymlvnkeeskpdnivytplvkelaqynvevhpytvrkdalpefftdvn qmydallnksgatgvftdfpdtgveflkgikS*AEHVLTMLNEHEVKFVDLRFTDTKGKEQHV TIPAHQVNAEFFEEGKMFDGSSIGGWKGINESDMVLMPDASTAVIDPFFADSTLIIRCDILE PGTLQGYDRDPRSIAKRAEDYLRATGIADTVLFGPEPEFFLFDDIRFGASISGSHVAIDDIE GAWNSSTKYEGGNKGHRPGVKGGYFPVPPVDSAQDIRSEMCLVMEQMGLVVEAHHHEVATAG QNEVATRFNTMTKKADEIQIYKYVVHNVAHRFGKTATFMPKPMFGDNGSGMHCHMSLAKNGT NLFSGDKYAGLSEQALYYIGGVIKHAKAINALANPTTNSYKRLVPGYEAPVMLAYSARNRSA SIRIPVVASPKARRIEVRFPDPAANPYLCFAALLMAGLDGIKNKIHPGEPMDKNLYDLPPEE AKEIPQVAGSLEEALNALDLDREFLKAGGVFTDEAIDAYIALRREEDDRVRMTPHPVEFELY YSV*
206	6H-3C-rTT-1f52	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAIHL VNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDSA*EHVLTMLNEHEVKFVDLRFTDTKGKEQHVTIPAHQVNAEFFE*

		EGKMFDGSSIGGWKGINESDMVLMPDASTAVIDPFFADSTLIIRCDILEPGTLQGYDRDPRS IAKRAEDYLRATGIADTVLFGPEPEFFLFDDIRFGASISGSHVAIDDIEGAWNSSTKYEGGN KGHRPGVKGGYFPVPPVDSAQDIRSEMCLVMEQMGLVVEAHHHEVATAGQNEVATRFNTMTK KADEIQIYKYVVHNVAHRFGKTATFMPKPMFGDNGSGMHCHMSLAKNGTNLFSGDKYAGLSE QALYYIGGVIKHAKAINALANPTTNSYKRLVPGYEAPVMLAYSARNRSASIRIPVVASPKAR RIEVRFPDPAANPYLCFAALLMAGLDGIKNKIHPGEPMDKNLYDLPPEEAKEIPQVAGSLEE ALNALDLDREFLKAGGVFTDEAIDAYIALRREEDDRVRMTPHPVEFELYYSV
207	6H-3C-rTT- ln4-1f52	MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAIHL VNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSI GSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYIN GVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSI TFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRLYNGL KFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIP LYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDK ILGCDWYFVPTDEGWTNDgggsSA*EHVLTMLNEHEVKFVDLRFTDTKGKEQHVTIPAHQVNA EFFEEGKMFDGSSIGGWKGINESDMVLMPDASTAVIDPFFADSTLIIRCDILEPGTLQGYDR DPRSIAKRAEDYLRATGIADTVLFGPEPEFFLFDDIRFGASISGSHVAIDDIEGAWNSSTKY EGGNKGHRPGVKGGYFPVPPVDSAQDIRSEMCLVMEQMGLVVEAHHHEVATAGQNEVATRFN TMTKKADEIQIYKYVVHNVAHRFGKTATFMPKPMFGDNGSGMHCHMSLAKNGTNLFSGDKYA GLSEQALYYIGGVIKHAKAINALANPTTNSYKRLVPGYEAPVMLAYSARNRSASIRIPVVAS PKARRIEVRFPDPAANPYLCFAALLMAGLDGIKNKIHPGEPMDKNLYDLPPEEAKEIPQVAG SLEEALNALDLDREFLKAGGVFTDEAIDAYIALRREEDDRVRMTPHPVEFELYYSV*
208	6H-3C-CRM_**degly-**1f52	GADDVVDSSKSFVMENFASYHGTKPGYVDSIQKGIQKPKSGTQGNYDDDWKEFYSTDNKYDA AGYSVDNENPLSGKAGGVVKVTYPGLTKVLALKVDNAETIKKELGLSLTEPLMEQVGTEEFI KRFGDGASRVVLSLPFAEGSSSVEYINNWEQAKALSVELEINFETRGKRGQDAMYEYMAQAC AGNRVRRSVGSSLSCINLDWDVIRDKTKTKIESLKEHGPIKNKMSESPNKAVSEEKAKQYLE EFHQTALEHPELSELKTVTGTNPVFAGANYAAWAVNVAQVIDSETADNLEKTTAALSILPGI GSVMGIADGAVHHNTEEIVAQSIALSSLMVAQAIPLVGELVDIGFAAYNFVESIINLFQVVH NSYNRPAYSPGHKTQPFLHDGYAVSWNTVEDSIIRTGFQGESGHDIKITAENTPLPIAGVLL PTIPGKLDVNKAKTHISVNGRKIRMRCRAIDGDVTFCRPKSPVYVGNGVHANLHVAFHRSSS EKIHSNEISSDSIGVLGYQKTVDHTKVNSKLSLFFEIKSS*AEHVLTMLNEHEVKFVDLRFTD TKGKEQHVTIPAHQVNAEFFEEGKMFDGSSIGGWKGINESDMVLMPDASTAVIDPFFADSTL IIRCDILEPGTLQGYDRDPRSIAKRAEDYLRATGIADTVLFGPEPEFFLFDDIRFGASISGS HVAIDDIEGAWNSSTKYEGGNKGHRPGVKGGYFPVPPVDSAQDIRSEMCLVMEQMGLVVEAH HHEVATAGQNEVATRFNTMTKKADEIQIYKYVVHNVAHRFGKTATFMPKPMFGDNGSGMHCH MSLAKNGTNLFSGDKYAGLSEQALYYIGGVIKHAKAINALANPTTNSYKRLVPGYEAPVMLA YSARNRSASIRIPVVASPKARRIEVRFPDPAANPYLCFAALLMAGLDGIKNKIHPGEPMDKN LYDLPPEEAKEIPQVAGSLEEALNALDLDREFLKAGGVFTDEAIDAYIALRREEDDRVRMTP HPVEFELYYSV*
209	6H-3C-CRM_**degly-ln4-**1f52	GADDVVDSSKSFVMENFASYHGTKPGYVDSIQKGIQKPKSGTQGNYDDDWKEFYSTDNKYDA AGYSVDNENPLSGKAGGVVKVTYPGLTKVLALKVDNAETIKKELGLSLTEPLMEQVGTEEFI KRFGDGASRVVLSLPFAEGSSSVEYINNWEQAKALSVELEINFETRGKRGQDAMYEYMAQAC AGNRVRRSVGSSLSCINLDWDVIRDKTKTKIESLKEHGPIKNKMSESPNKAVSEEKAKQYLE EFHQTALEHPELSELKTVTGTNPVFAGANYAAWAVNVAQVIDSETADNLEKTTAALSILPGI GSVMGIADGAVHHNTEEIVAQSIALSSLMVAQAIPLVGELVDIGFAAYNFVESIINLFQVVH NSYNRPAYSPGHKTQPFLHDGYAVSWNTVEDSIIRTGFQGESGHDIKITAENTPLPIAGVLL PTIPGKLDVNKAKTHISVNGRKIRMRCRAIDGDVTFCRPKSPVYVGNGVHANLHVAFHRSSS EKIHSNEISSDSIGVLGYQKTVDHTKVNSKLSLFFEIKSgggsSA*EHVLTMLNEHEVKFVDL RFTDTKGKEQHVTIPAHQVNAEFFEEGKMFDGSSIGGWKGINESDMVLMPDASTAVIDPFFA DSTLIIRCDILEPGTLQGYDRDPRSIAKRAEDYLRATGIADTVLFGPEPEFFLFDDIRFGAS ISGSHVAIDDIEGAWNSSTKYEGGNKGHRPGVKGGYFPVPPVDSAQDIRSEMCLVMEQMGLV VEAHHHEVATAGQNEVATRFNTMTKKADEIQIYKYVVHNVAHRFGKTATFMPKPMFGDNGSG MHCHMSLAKNGTNLFSGDKYAGLSEQALYYIGGVIKHAKAINALANPTTNSYKRLVPGYEAP VMLAYSARNRSASIRIPVVASPKARRIEVRFPDPAANPYLCFAALLMAGLDGIKNKIHPGEP MDKNLYDLPPEEAKEIPQVAGSLEEALNALDLDREFLKAGGVFTDEAIDAYIALRREEDDRV RMTPHPVEFELYYSV*
210	6H-3C-CRM_**degly-ln8-**1f52	GADDVVDSSKSFVMENFASYHGTKPGYVDSIQKGIQKPKSGTQGNYDDDWKEFYSTDNKYDA AGYSVDNENPLSGKAGGVVKVTYPGLTKVLALKVDNAETIKKELGLSLTEPLMEQVGTEEFI KRFGDGASRVVLSLPFAEGSSSVEYINNWEQAKALSVELEINFETRGKRGQDAMYEYMAQAC AGNRVRRSVGSSLSCINLDWDVIRDKTKTKIESLKEHGPIKNKMSESPNKAVSEEKAKQYLE EFHQTALEHPELSELKTVTGTNPVFAGANYAAWAVNVAQVIDSETADNLEKTTAALSILPGI GSVMGIADGAVHHNTEEIVAQSIALSSLMVAQAIPLVGELVDIGFAAYNFVESIINLFQVVH

		NSYNRPAYSPGHKTQPFLHDGYAVSWNTVEDSIIRTGFQGESGHDIKITAENTPLPIAGVLL PTIPGKLDVNKAKTHISVNGRKIRMRCRAIDGDVTFCRPKSPVYVGNGVHANLHVAFHRSSS EKIHSNEISSDSIGVLGYQKTVDHTKVNSKLSLFFEIKSgggsgggsSA*EHVLTMLNEHEVK FVDLRFTDTKGKEQHVTIPAHQVNAEFFEEGKMFDGSSIGGWKGINESDMVLMPDASTAVID PFFADSTLIIRCDILEPGTLQGYDRDPRSIAKRAEDYLRATGIADTVLFGPEPEFFLFDDIR FGASISGSHVAIDDIEGAWNSSTKYEGGNKGHRPGVKGGYFPVPPVDSAQDIRSEMCLVMEQ MGLVVEAHHHEVATAGQNEVATRFNTMTKKADEIQIYKYVVHNVAHRFGKTATFMPKPMFGD NGSGMHCHMSLAKNGTNLFSGDKYAGLSEQALYYIGGVIKHAKAINALANPTTNSYKRLVPG YEAPVMLAYSARNRSASIRIPVVASPKARRIEVRFPDPAANPYLCFAALLMAGLDGIKNKIH PGEPMDKNLYDLPPEEAKEIPQVAGSLEEALNALDLDREFLKAGGVFTDEAIDAYIALRREE DDRVRMTPHPVEFELYYSV*
HIV capsid oligerization domain (HIV)		
211	HIV-CA-3P0A-rTT	*PIVQNLQGQMVHQAISCLCLNAWVKVVEEKAFSPEVIPMFSALSEGATPQDLNTMLNTVGGH QAAMQMLKETINEEAAEWDRLHPVHAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMTHNPPI PVGEIYKRWIILGLNKIVRMYSPTSILDIRQGPKEPFRDYVDRFYKTLRAEQASQEVKNAAT ETLLVQNANPDCKTILKALGPGATLEEMMTACQGVGGPGHKARV*KNLDCWVDNEEDIDVILK KSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAIHLVNNESSEVIVHKAMDIEYN DMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKD SAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGVLMGSAEITGLGAIREDN NITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITFLRDFWGNPLRYDTEYYL IPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNEIDSFVKS GDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQ LKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND
212	HIV-CA-3P0A-rTT-858	*PIVQNLQGQMVHQAISCLCLNAWVKVVEEKAFSPEVIPMFSALSEGATPQDLNTMLNTVGGH QAAMQMLKETINEEAAEWDRLHPVHAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMTHNPPI PVGEIYKRWIILGLNKIVRMYSPTSILDIRQGPKEPFRDYVDRFYKTLRAEQASQEVKNAAT ETLLVQNANPDCKTILKALGPGATLEEMMTACQGVGGPGHKARV*pipfsysKNLDCWVDNEE DIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAIHLVNNESSEVIVHK AMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNN LIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGVLMGSAEITGL GAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITFLRDFWGNPLR YDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNE IDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRD LKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFVPTD EGWTND
213	HIV-CA-3P0A-rTT-836	*PIVQNLQGQMVHQAISCLCLNAWVKVVEEKAFSPEVIPMFSALSEGATPQDLNTMLNTVGGH QAAMQMLKETINEEAAEWDRLHPVHAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMTHNPPI PVGEIYKRWIILGLNKIVRMYSPTSILDIRQGPKEPFRDYVDRFYKTLRAEQASQEVKNAAT ETLLVQNANPDCKTILKALGPGATLEEMMTACQGVGGPGHKARV*skfigitelkkleskink vfsTpipfsysKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLV PGINGKAIHLVNNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSII SSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITND RLSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEI EKLYTSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLN IYYRRLYNGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRIL RVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIAS NWYFNHLKDKILGCDWYFVPTDEGWTND
214	HIV-CA-3P0A-rTT-217-839	*PIVQNLQGQMVHQAISCLCLNAWVKVVEEKAFSPEVIPMFSALSEGATPQDLNTMLNTVGGH QAAMQMLKETINEEAAEWDRLHPVHAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMTHNPPI PVGEIYKRWIILGLNKIVRMYSPTSILDIRQGPKEPFRDYVDRFYKTLRAEQASQEVKNAAT ETLLVQNANPDCKTILKALGPGATLEEMMTA*igitelkkleskinkvfsTpipfsysKNLDC WVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAIHLVNNESS EVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSV SLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGVLMGS AEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITFLRDF WGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKR YTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKME AVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDW YFVPTDEGWTND

215	HIV-CA-3P0A-rTT-217-840	*PIVQNLQGQMVHQAISCLCLNAWVKVVEEKAFSPEVIPMFSALSEGATPQDLNTMLNTVGGH QAAMQMLKETINEEAAEWDRLHPVHAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMTHNPPI PVGEIYKRWIILGLNKIVRMYSPTSILDIRQGPKEPFRDYVDRFYKTLRAEQASQEVKNAAT ETLLVQNANPDCKTILKALGPGATLEEMMTA*gitelkkleskinkvfsTpipfsysKNLDCW VDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAIHLVNNESSE VIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVS LKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGVLMGSA EITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITFLRDFW GNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRY TPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEA VKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWY FVPTDEGWTND
216	HIV-CA-3P0A-rTT-217-841	*PIVQNLQGQMVHQAISCLCLNAWVKVVEEKAFSPEVIPMFSALSEGATPQDLNTMLNTVGGH QAAMQMLKETINEEAAEWDRLHPVHAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMTHNPPI PVGEIYKRWIILGLNKIVRMYSPTSILDIRQGPKEPFRDYVDRFYKTLRAEQASQEVKNAAT ETLLVQNANPDCKTILKALGPGATLEEMMTA*itelkkleskinkvfsTpipfsysKNLDCWV DNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAIHLVNNESSEV IVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSL KGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGVLMGSAE ITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITFLRDFWG NPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYT PNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAV KLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYF VPTDEGWTND
217	HIV-CA-3P0A-rTT-217-842	*PIVQNLQGQMVHQAISCLCLNAWVKVVEEKAFSPEVIPMFSALSEGATPQDLNTMLNTVGGH QAAMQMLKETINEEAAEWDRLHPVHAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMTHNPPI PVGEIYKRWIILGLNKIVRMYSPTSILDIRQGPKEPFRDYVDRFYKTLRAEQASQEVKNAAT ETLLVQNANPDCKTILKALGPGATLEEMMTA*telkkleskinkvfsTpipfsysKNLDCWVD NEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAIHLVNNESSEVI VHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLK GNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGVLMGSAEI TGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITFLRDFWGN PLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTP NNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVK LRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFV PTDEGWTND
218	HIV-CA-3P0A-rTT-217-843	*PIVQNLQGQMVHQAISCLCLNAWVKVVEEKAFSPEVIPMFSALSEGATPQDLNTMLNTVGGH QAAMQMLKETINEEAAEWDRLHPVHAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMTHNPPI PVGEIYKRWIILGLNKIVRMYSPTSILDIRQGPKEPFRDYVDRFYKTLRAEQASQEVKNAAT ETLLVQNANPDCKTILKALGPGATLEEMMTA*telkkleskinkvfsTpipfsysKNLDCWVD NEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAIHLVNNESSEVI VHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLK GNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGVLMGSAEI TGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITFLRDFWGN PLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTP NNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVK LRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFV PTDEGWTND
219	HIV-CA-3H4E-rTT	*PIVQNLQGQMVHQCISPRTLNAWVKVVEEKAFSPEVIPMFSALSCGATPQDLNTMLNTVGGH QAAMQMLKETINEEAAEWDRLHPVHAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMTHNPPI PVGEIYKRWIILGLNKIVRMYSPTSILDIRQGPKEPFRDYVDRFYKTLRAEQASQEVKNAAT ETLLVQNANPDCKTILKALGPGATLEEMMTACQGVGGPGHKARVLK*NLDCWVDNEEDIDVIL KKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAIHLVNNESSEVIVHKAMDIEY NDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLK DSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYINGVLMGSAEITGLGAIRED NNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSITFLRDFWGNPLRYDTEYY LIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNEIDSFVK SGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDLKTYSV QLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND

Encapsulin		
254	EN-TThc-degly	MEFLKRSFAPLTEKQWQEIDNRAREIFKTQLYGRKFVDVEGPYGWEYAAHPLCEVEVLSDEN EVVKWGLRKSLPLIELRATFTLDLWELDNLECGKPNVDLSSLEETVRKVAEFEDEVIFRGCE KSGVKGLLSFEERKIECGSTPKDLLEAIVRALSIFSKDGIEGPYTLVINTDRWINFLKEEAG HYPLEKRVEECLRGGKIITTPRIEDALVVSERGGDFKLILGQDLSIGYEDREKDAVRLFITE TFTFQVVNPEALILLKFgsgsgsMKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGF NSAVITYPDAQLVPGINGKAIHLVNNEASEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSAS HLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAY LANKWVFITITNDRLSSANLYINGVLMGSAEITGLGAIREDNQITLKLDRCNNNNQYVSIDK FRIFCKALNPKEIEKLYTSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKQITDYMY LTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYP KDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKQASLGLVGTHNGQ IGNDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTND
255	EN-CRM197-degly	MEFLKRSFAPLTEKQWQEIDNRAREIFKTQLYGRKFVDVEGPYGWEYAAHPLCEVEVLSDEN EVVKWGLRKSLPLIELRATFTLDLWELDNLECGKPNVDLSSLEETVRKVAEFEDEVIFRGCE KSGVKGLLSFEERKIECGSTPKDLLEAIVRALSIFSKDGIEGPYTLVINTDRWINFLKEEAG HYPLEKRVEECLRGGKIITTPRIEDALVVSERGGDFKLILGQDLSIGYEDREKDAVRLFITE TFTFQVVNPEALILLKFgsgsgsGADDVVDSSKSFVMENFASYHGTKPGYVDSIQKGIQKPK SGTQGNYDDDWKEFYSTDNKYDAAGYSVDNENPLSGKAGGVVKVTYPGLTKVLALKVDNAET IKKELGLSLTEPLMEQVGTEEFIKRFGDGASRVVLSLPFAEGSSSVEYINNWEQAKALSVEL EINFETRGKRGQDAMYEYMAQACAGNRVRRSVGSSLSCINLDWDVIRDKTKTKIESLKEHGP IKNKMSESPNKAVSEEKAKQYLEEFHQTALEHPELSELKTVTGTNPVFAGANYAAWAVNVAQ VIDSETADNLEKTTAALSILPGIGSVMGIADGAVHHNTEEIVAQSIALSSLMVAQAIPLVGE LVDIGFAAYNFVESIINLFQVVHNSYNRPAYSPGHKTQPFLHDGYAVSWNTVEDSIIRTGFQ GESGHDIKITAENTPLPIAGVLLPTIPGKLDVNKAKTHISVNGRKIRMRCRAIDGDVTFCRP KSPVYVGNGVHANLHVAFHRSSSEKIHSNEISSDSIGVLGYQKTVDHTKVNSKLSLFFEIKS
256	EN-2xCRM197-degly	MEFLKRSFAPLTEKQWQEIDNRAREIFKTQLYGRKFVDVEGPYGWEYAAHPLCEVEVLSDEN EVVKWGLRKSLPLIELRATFTLDLWELDNLECGKPNVDLSSLEETVRKVAEFEDEVIFRGCE KSGVKGLLSFEERKIECGSTPKDLLEAIVRALSIFSKDGIEGPYTLVINTDRWINFLKEEAG HYPLEKRVEECLRGGKIITTPRIEDALVVSERGGDFKLILGQDLSIGYEDREKDAVRLFITE TFTFQVVNPEALILLKFgsgDYKDDDDKgsgGADDVVDSSKSFVMENFASYHGTKPGYVDSI QKGIQKPKSGTQGNYDDDWKEFYSTDNKYDAAGYSVDNENPLSGKAGGVVKVTYPGLTKVLA LKVDNAETIKKELGLSLTEPLMEQVGTEEFIKRFGDGASRVVLSLPFAEGSSSVEYINNWEQ AKALSVELEINFETRGKRGQDAMYEYMAQACAGNRVRRSVGSSLSCINLDWDVIRDKTKTKI ESLKEHGPIKNKMSESPNKAVSEEKAKQYLEEFHQTALEHPELSELKTVTGTNPVFAGANYA AWAVNVAQVIDSETADNLEKTTAALSILPGIGSVMGIADGAVHHNTEEIVAQSIALSSLMVA QAIPLVGELVDIGFAAYNFVESIINLFQVVHNSYNRPAYSPGHKTQPFLHDGYAVSWNTVED SIIRTGFQGESGHDIKITAENTPLPIAGVLLPTIPGKLDVNKAKTHISVNGRKIRMRCRAID GDVTFCRPKSPVYVGNGVHANLHVAFHRSSSEKIHSNEISSDSIGVLGYQKTVDHTKVNSKL SLFFEIKSgggsgggsGADDVVDSSKSFVMENFASYHGTKPGYVDSIQKGIQKPKSGTQGNY DDDWKEFYSTDNKYDAAGYSVDNENPLSGKAGGVVKVTYPGLTKVLALKVDNAETIKKELGL SLTEPLMEQVGTEEFIKRFGDGASRVVLSLPFAEGSSSVEYINNWEQAKALSVELEINFETR GKRGQDAMYEYMAQACAGNRVRRSVGSSLSCINLDWDVIRDKTKTKIESLKEHGPIKNKMSE SPNKAVSEEKAKQYLEEFHQTALEHPELSELKTVTGTNPVFAGANYAAWAVNVAQVIDSETA DNLEKTTAALSILPGIGSVMGIADGAVHHNTEEIVAQSIALSSLMVAQAIPLVGELVDIGFA AYNFVESIINLFQVVHNSYNRPAYSPGHKTQPFLHDGYAVSWNTVEDSIIRTGFQGESGHDI KITAENTPLPIAGVLLPTIPGKLDVNKAKTHISVNGRKIRMRCRAIDGDVTFCRPKSPVYVG NGVHANLHVAFHRSSSEKIHSNEISSDSIGVLGYQKTVDHTKVNSKLSLFFEIKS
257	EN-HID	MEFLKRSFAPLTEKQWQEIDNRAREIFKTQLYGRKFVDVEGPYGWEYAAHPLCEVEVLSDEN EVVKWGLRKSLPLIELRATFTLDLWELDNLECGKPNVDLSSLEETVRKVAEFEDEVIFRGCE KSGVKGLLSFEERKIECGSTPKDLLEAIVRALSIFSKDGIEGPYTLVINTDRWINFLKEEAG HYPLEKRVEECLRGGKIITTPRIEDALVVSERGGDFKLILGQDLSIGYEDREKDAVRLFITE TFTFQVVNPEALILLKFgsgDYKDDDDKgsgsnmantqmksdkiiiahrgasgylpehtles kalafaqqadyleqdlamtkdgrlvvihdhfldgltdvakkfphrhrkdgryyvidftlkei qslemtenfetkdgkqaqvypnrfplwkshfrihtfedeiefiqglekstgkkvgiypeika pwfhhqngkdiaaetlkvlkkygydkktdmvylqtfdfnelkriktellpqmgmdlklvqli aytdwketqekdpkgywvnynydwmfkpgamaevvkyadgvgpgwymlvnkeeskpdnivyt plvkelaqynvevhpytvrkdalpefftdvnqmydallnksgatgvftdfpdtgveflkgik
369	EN - glyser-	MEFLKRSFAPLTEKQWQEIDNRAREIFKTQLYGRKFVDVEGPYGWEYAAHPLCEVEVLSDEN EVVKWGLRKSLPLIELRATFTLDLWELDNLECGKPNVDLSSLEETVRKVAEFEDEVIFRGCE KSGVKGLLSFEERKIECGSTPKDLLEAIVRALSIFSKDGIEGPYTLVINTDRWINFLKEEAG

	G53C/K96C -rTT	HYPLEKRVEECLRGGKIITTPRIEDALVVSERGGDFKLILGQDLSIGYEDREKDAVRLFITE TFTFQVVNPEALILLKFgggsgggsgggsKnldcwvdneedidvilkkstilnldinndiis disgfnssvitypdaqlvpgingkaihlvnnessevivhkamdieyndmfnnftvsfwlrvp kvsashleqygtneysiissmkkhslsigsgwsvslkgnnliwtlkdsagevrqitfrdlpd kfnaylankwvfititndrlssanlyingvlmgsaeitglgairednnitlkldrcnnnnqy vsidkfrifckalnpkeieklytsylsitflrdfwgnplrydteyylipvassskdvqlkni tdymyltnapsytngklniyyrrlynglkfiikrytpnneidsfvksgdfiklyvsynnneh ivgypkdgnafnnldrilrvgynapgiplykkmeavklrdlktysvqlklyddknaslglvg thngqignpnrdiliasnwyfnhlkdkilgcdwyfvptdegwtnd
370	EN - glyser-G53C/K96C - rTT	MEFLKRSFAPLTEKQWQEIDNRAREIFKTQLYGRKFVDVEGPYGWEYAAHPLCEVEVLSDEN EVVKWGLRKSLPLIELRATFTLDLWELDNLECGKPNVDLSSLEETVRKVAEFEDEVIFRGCE KSGVKGLLSFEERKIECGSTPKDLLEAIVRALSIFSKDGIEGPYTLVINTDRWINFLKEEAG HYPLEKRVEECLRGGKIITTPRIEDALVVSERGGDFKLILGQDLSIGYEDREKDAVRLFITE TFTFQVVNPEALILLKFgggsgggsgggslrdfwgnplrydteyylipvassskdvqlknit dymyltnapsytngklniyyrrlynglkfiikrytpnneidsfvksgdfiklyvsynnnehi vgypkdgnafnnldrilrvgynapgiplykkmeavklrdlktysvqlklyddknaslglvgt hngqignpnrdiliasnwyfnhlkdkilgcdwyfvptdegwtnd
371	EN - glyser-G53C/K96C - rTT N249	MEFLKRSFAPLTEKQWQEIDNRAREIFKTQLYGRKFVDVEGPYGWEYAAHPLCEVEVLSDEN EVVKWGLRKSLPLIELRATFTLDLWELDNLECGKPNVDLSSLEETVRKVAEFEDEVIFRGCE KSGVKGLLSFEERKIECGSTPKDLLEAIVRALSIFSKDGIEGPYTLVINTDRWINFLKEEAG HYPLEKRVEECLRGGKIITTPRIEDALVVSERGGDFKLILGQDLSIGYEDREKDAVRLFITE TFTFQVVNPEALILLKFgggsgggsgggs knldcwvdneedidvilkkstilnldinndiisdisgfnssvitypdaqlvpgingkaihlv nnessevivhkamdieyndmfnnftvsfwlrvpkvsashleqygtneysiissmkkhslsig sgwsvslkgnnliwtlkdsagevrqitfrdlpdkfnaylankwvfititndrlssanlying vlmgsaeitglgairednnitlkldrcnnnnqyvsidkfrifckalnpkeieklytsylsit
372	EN - glyser-G53C/K96C - rTT N193	MEFLKRSFAPLTEKQWQEIDNRAREIFKTQLYGRKFVDVEGPYGWEYAAHPLCEVEVLSDEN EVVKWGLRKSLPLIELRATFTLDLWELDNLECGKPNVDLSSLEETVRKVAEFEDEVIFRGCE KSGVKGLLSFEERKIECGSTPKDLLEAIVRALSIFSKDGIEGPYTLVINTDRWINFLKEEAG HYPLEKRVEECLRGGKIITTPRIEDALVVSERGGDFKLILGQDLSIGYEDREKDAVRLFITE TFTFQVVNPEALILLKFgggsgggsgggsknldcwvdneedidvilkkstilnldinndiis disgfnssvitypdaqlvpgingkaihlvnnessevivhkamdieyndmfnnftvsfwlrvp kvsashleqygtneysiissmkkhslsigsgwsvslkgnnliwtlkdsagevrqitfrdlpd kfnaylankwvfititndrlssanlyingvlmgsae
373	EN - glyser-G53C/K96C - rTT N87	MEFLKRSFAPLTEKQWQEIDNRAREIFKTQLYGRKFVDVEGPYGWEYAAHPLCEVEVLSDEN EVVKWGLRKSLPLIELRATFTLDLWELDNLECGKPNVDLSSLEETVRKVAEFEDEVIFRGCE KSGVKGLLSFEERKIECGSTPKDLLEAIVRALSIFSKDGIEGPYTLVINTDRWINFLKEEAG HYPLEKRVEECLRGGKIITTPRIEDALVVSERGGDFKLILGQDLSIGYEDREKDAVRLFITE TFTFQVVNPEALILLKFgggsgggsgggsKnldcwvdneedidvilkkstilnldinndiis disgfnssvitypdaqlvpgingkaihlvnnessevivhkamdieyndmfnnf
374	EN - glyser-K146C/A185C - rTT N87	MEFLKRSFAPLTEKQWQEIDNRAREIFKTQLYGRKFVDVEGPYGWEYAAHPLCEVEVLSDEN EVVKWGLRKSLPLIELRATFTLDLWELDNLErGcPNVDLSSLEETVRKVAEFEDEVIFRGCE KSGVKGLLSFEERKIECGSTPKDLLEAIVRALSIFSKDGIEGPYTLVINTDRWINFLKEEAG HYPLEKRVEECLRGGKIITTPRIEDALVVSERGGDFKLILGQDLSIGYEDREKDAVRLFITE TFTFQVVNPEALILLKFgggsgggsgggsKnldcwvdneedidvilkkstilnldinndiis disgfnssvitypdaqlvpgingkaihlvnnessevivhkamdieyndmfnnf
375	EN - glyser-G53C/K96C rTT N87	MEFLKRSFAPLTEKQWQEIDNRAREIFKTQLYGRKFVDVEGPYGWEYAAHPLgEVEVLSDEN EVVKWGLRKSLPLIELRATFTLDLWELDNLErGkPNVDLSSLEETVRKVAEFEDEVIFRGCE KSGVKGLLSFEERKIECGSTPcDLLEAIVRALSIFSKDGIEGPYTLVINTDRWINFLKEEcG HYPLEKRVEECLRGGKIITTPRIEDALVVSERGGDFKLILGQDLSIGYEDREKDAVRLFITE TFTFQVVNPEALILLKFgggsgggsgggsKnldcwvdneedidvilkkstilnldinndiis disgfnssvitypdaqlvpgingkaihlvnnessevivhkamdieyndmfnnf
376	EN - caIgG2a-G53C/K96C - rTT N88	MEFLKRSFAPLTEKQWQEIDNRAREIFKTQLYGRKFVDVEGPYGWEYAAHPLCEVEVLSDEN EVVKWGLRKSLPLIELRATFTLDLWELDNLECGKPNVDLSSLEETVRKVAEFEDEVIFRGCE KSGVKGLLSFEERKIECGSTPKDLLEAIVRALSIFSKDGIEGPYTLVINTDRWINFLKEEAG HYPLEKRVEECLRGGKIITTPRIEDALVVSERGGDFKLILGQDLSIGYEDREKDAVRLFITE TFTFQVVNPEALILLKFggEPKIPQPQPKPQPQPQPKPQPKPEPEggKnldcwvdneedi dvilkkstilnldinndiisdisgfnssvitypdaqlvpgingkaihlvnnessevivhkam dieyndmfnnf
377	EN -CD8-G53C/K96C -rTT N88	MEFLKRSFAPLTEKQWQEIDNRAREIFKTQLYGRKFVDVEGPYGWEYAAHPLCEVEVLSDEN EVVKWGLRKSLPLIELRATFTLDLWELDNLECGKPNVDLSSLEETVRKVAEFEDEVIFRGCE KSGVKGLLSFEERKIECGSTPKDLLEAIVRALSIFSKDGIEGPYTLVINTDRWINFLKEEAG

		HYPLEKRVEECLRGGKIITTPRIEDALVVSERGGDFKLILGQDLSIGYEDREKDAVRLFITE TFTFQVVNPEALILLKFggKPTTTPAPRPPTPAPTIASQPLSLRPEACRPAAGGAVHTRGLD FACDggKnldcwvdneedidvilkkstilnldinndiisdisgfnssvitypdaqlvpging kaihlvnnessevivhkamdieyndmfnnf
378	EN - hinge-G53C/K96C -rTT N88	MEFLKRSFAPLTEKQWQEIDNRAREIFKTQLYGRKFVDVEGPYGWEYAAHPLCEVEVLSDEN EVVKWGLRKSLPLIELRATFTLDLWELDNLECGKPNVDLSSLEETVRKVAEFEDEVIFRGCE KSGVKGLLSFEERKIECGSTPKDLLEAIVRALSIFSKDGIEGPYTLVINTDRWINFLKEEAG HYPLEKRVEECLRGGKIITTPRIEDALVVSERGGDFKLILGQDLSIGYEDREKDAVRLFITE TFTFQVVNPEALILLKFggEPKSDKTHTPPPAPELLgsgEPKSDKTHTPPPAPELLggKnld cwvdneedidvilkkstilnldinndiisdisgfnssvitypdaqlvpgingkaihlvnnes sevivhkamdieyndmfnnf
HBV		
379	rTT-hinge-HBV P25C/R127C	KNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAIHLV NNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIG SGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYING VLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSIT FLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRLYNGLK FIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPL YKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKI LGCDWYFVPTDEGWTNDggsgEPKSDKTHTPPPAPELLgsgEPKSDKTHTPPPAPELLgsgg MDIDPYKEFGATVELLSFLPSDFFcSVRDLLDTASALYREALESPEHCSPHHTALRQAILCW GELMTLATWVGVNLEDPASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETVIEYLVSFGV WIcTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRRSQSRESQC
380	rTT-hinge-HBV E14C/A36C	KNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAIHLV NNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIG SGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYING VLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSIT FLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRLYNGLK FIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPL YKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKI LGCDWYFVPTDEGWTNDggsgEPKSDKTHTPPPAPELLgsgEPKSDKTHTPPPAPELLgsgg MDIDPYKEFGATVcLLSFLPSDFFPSVRDLLDTAScLYREALESPEHCSPHHTALRQAILCW GELMTLATWVGVNLEDPASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETVIEYLVSFGV WIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRRSQSRESQC
381	rTT-hinge-HBV D29C/R127C	KNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAIHLV NNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIG SGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYING VLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSIT FLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRLYNGLK FIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPL YKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKI LGCDWYFVPTDEGWTNDggsgEPKSDKTHTPPPAPELLgsgEPKSDKTHTPPPAPELLgsgg MDIDPYKEFGATVELLSFLPSDFFPSVRcLLDTASALYREALESPEHCSPHHTALRQAILCW GELMTLATWVGVNLEDPASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETVIEYLVSFGV WIcTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRRSQSRESQC
382	rTT-caIgG2a-HBV P25C/R127C	KNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAIHLV NNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIG SGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYING VLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSIT FLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRLYNGLK FIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPL YKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKI LGCDWYFVPTDEGWTNDgsgGSEPKIPQPQPKPQPQPQPQPKPQPKPEPEgsgMDIDPYKEF GATVELLSFLPSDFFcSVRDLLDTASALYREALESPEHCSPHHTALRQAILCWGELMTLATW VGVNLEDPASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETVIEYLVSFGVWIcTPPAYR PPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRRSQSRESQC
383	rTT-caIgG2a-HBV E14C/A36C	KNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAIHLV NNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIG SGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYING VLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSIT FLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRLYNGLK

		FIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPL YKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKI LGCDWYFVPTDEGWTNDgsgGSEPKIPQPQPKPQPQPQPQPKPQPKPEPEgsgMDIDPYKEF GATVcLLSFLPSDFFPSVRDLLDTAScLYREALESPEHCSPHHTALRQAILCWGELMTLATW VGVNLEDPASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETVIEYLVSFGVWIRTPPAYR PPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRRSQSRESQC
384	rTT-caIgG2a-HBV D29C/R127C	KNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAIHLV NNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIG SGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYING VLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSIT FLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRLYNGLK FIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPL YKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKI LGCDWYFVPTDEGWTNDgsgGSEPKIPQPQPKPQPQPQPQPKPQPKPEPEgsgMDIDPYKEF GATVELLSFLPSDFFPSVRcLLDTASALYREALESPEHCSPHHTALRQAILCWGELMTLATW VGVNLEDPASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRETVIEYLVSFGVWIcTPPAYR PPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRRSQSRESQC
385	rTT-CD8-HBV P25C/R127C	KNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAIHLV NNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIG SGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYING VLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSIT FLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRLYNGLK FIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPL YKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKI LGCDWYFVPTDEGWTNDgsggsgKPTTTPAPRPPTPAPTIASQPLSLRPEACRPAAGGAVHT RGLDFACDgsgMDIDPYKEFGATVELLSFLPSDFFcSVRDLLDTASALYREALESPEHCSPH HTALRQAILCWGELMTLATWVGVNLEDPASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRE TVIEYLVSFGVWIcTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRR SQSRESQC
386	rTT-CD8-HBV E14C/A36C	KNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAIHLV NNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIG SGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYING VLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSIT FLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRLYNGLK FIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPL YKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKI LGCDWYFVPTDEGWTNDgsggsgKPTTTPAPRPPTPAPTIASQPLSLRPEACRPAAGGAVHT RGLDFACDgsgMDIDPYKEFGATVcLLSFLPSDFFPSVRDLLDTAScLYREALESPEHCSPH HTALRQAILCWGELMTLATWVGVNLEDPASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRE TVIEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRR SQSRESQC
387	rTT-CD8-HBV D29C/R127C	KNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAIHLV NNESSEVIVHKAMDIEYNDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIG SGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKFNAYLANKWVFITITNDRLSSANLYING VLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEKLYTSYLSIT FLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRLYNGLK FIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPL YKKMEAVKLRDLKTYSVQLKLYDDKNASLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKI LGCDWYFVPTDEGWTNDgsggsgKPTTTPAPRPPTPAPTIASQPLSLRPEACRPAAGGAVHT RGLDFACDgsgMDIDPYKEFGATVELLSFLPSDFFPSVRcLLDTASALYREALESPEHCSPH HTALRQAILCWGELMTLATWVGVNLEDPASRDLVVSYVNTNMGLKFRQLLWFHISCLTFGRE TVIEYLVSFGVWIcTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRR SQSRESQC
AP205		
388	AP205-glyser-rTT	MANKPMQPITSTANKIVWSDPTRLSTTFSASLLRQRVKVGIAELNNVSGQYVSVYKRPAPKP EGCADACVIMPNENQSIRTVISGSAENLATLKAEWETHKRNVDTLFASGNAGLGFLDPTAAI VSSDTTgggsgggsgggsKnldcwvdneedidvilkkstilnldinndiisdisgfnssvit ypdaqlvpgingkaihlvnnessevivhkamdieyndmfnnftvsfwlrvpkvsashleqyg tneysiissmkkhslsigsgwsvslkgnnliwtlkdsagevrqitfrdlpdkfnaylankwv fititndrlssanlyingvlmgsaeitglgairednnitlkldrcnnnnqyvsidkfrifck alnpkeieklytsylsitflrdfwgnplrydteyylipvassskdvqlknitdymyltnaps

		ytngklniyyrrlynglkfiikrytpnneidsfvksgdfiklyvsynnnehivgypkdgnaf nnldrilrvgynapgiplykkmeavklrdlktysvqlklyddknaslglvgthngqignd pn rdiliasnwyfnhlkdkilgcdwyfvptdegwtnd
389	AP205-glyser-rTT N193	MANKPMQPITSTANKIVWSDPTRLSTTFSASLLRQRVKVGIAELNNVSGQYVSVYKRPAPKP EGCADACVIMPNENQSIRTVISGSAENLATLKAEWETHKRNVDTLFASGNAGLGFLDPTAAI VSSDTTgggsgggsgggsknldcwvdneedidvilkkstilnldinndiisdisgfnssvit ypdaqlvpgingkaihlvnnessevivhkamdieyndmfnnftvsfwlrvpkvsashleqyg tneysiissmkkhslsigsgwsvslkgnnliwtlkdsagevrqitfrdlpdkfnaylankwv fititndrlssanlyingvlmgsae
390	AP205-glyser-rTT N87	MANKPMQPITSTANKIVWSDPTRLSTTFSASLLRQRVKVGIAELNNVSGQYVSVYKRPAPKP EGCADACVIMPNENQSIRTVISGSAENLATLKAEWETHKRNVDTLFASGNAGLGFLDPTAAI VSSDTTgggsgggsgggsKnldcwvdneedidvilkkstilnldinndiisdisgfnssvit ypdaqlvpgingkaihlvnnessevivhkamdieyndmfnnf
391	AP205-caIgG2a-rTT N87	MANKPMQPITSTANKIVWSDPTRLSTTFSASLLRQRVKVGIAELNNVSGQYVSVYKRPAPKP EGCADACVIMPNENQSIRTVISGSAENLATLKAEWETHKRNVDTLFASGNAGLGFLDPTAAI VSSDTTgsgEPKIPQPQPKPQPQPQPQPKPQPKPEPEgsgKnldcwvdneedidvilkksti lnldinndiisdisgfnssvitypdaqlvpgingkaihlvnnessevivhkamdieyndmfn nf
392	AP205-CD8-rTT N87	MANKPMQPITSTANKIVWSDPTRLSTTFSASLLRQRVKVGIAELNNVSGQYVSVYKRPAPKP EGCADACVIMPNENQSIRTVISGSAENLATLKAEWETHKRNVDTLFASGNAGLGFLDPTAAI VSSDTTgsgKPTTT PAPRPPTPAPTIASQPLSLRPEACRPAAGGAVHTRGLDFACDgsgKnldcwvdneedidvil kkstilnldinndiisdisgfnssvitypdaqlvpgingkaihlvnnessevivhkamdiey ndmfnnf
393	AP205-hinge-rTT N87	MANKPMQPITSTANKIVWSDPTRLSTTFSASLLRQRVKVGIAELNNVSGQYVSVYKRPAPKP EGCADACVIMPNENQSIRTVISGSAENLATLKAEWETHKRNVDTLFASGNAGLGFLDPTAAI VSSDTTgsgEPKSDKTHTPPPAPELLgsgEPKSDKTHTPPPAPELLgsgKnldcwvdneedi dvilkkstilnldinndiisdisgfnssvitypdaqlvpgingkaihlvnnessevivhkam dieyndmfnnf
394	AP205 T81C-hinge-rTT N87	MANKPMQPITSTANKIVWSDPTRLSTTFSASLLRQRVKVGIAELNNVSGQYVSVYKRPAPKP EGCADACVIMPNENQSIRcVISGSAENLATLKAEWETHKRNVDTLFASGNAGLGFLDPTAAI VSSDTTgsgEPKSDKTHTPPPAPELLgsgEPKSDKTHTPPPAPELLgsgKnldcwvdneedi dvilkkstilnldinndiisdisgfnssvitypdaqlvpgingkaihlvnnessevivhkam dieyndmfnnf
395	AP205 S53C/H100C- hinge-rTT N87	MANKPMQPITSTANKIVWSDPTRLSTTFSASLLRQRVKVGIAELNNVSGQYVcVYKRPAPKP EGCADACVIMPNENQSIRTVISGSAENLATLKAEWETcKRNVDTLFASGNAGLGFLDPTAAI VSSDTTgsgEPKSDKTHTPPPAPELLgsgEPKSDKTHTPPPAPELLgsgKnldcwvdneedi dvilkkstilnldinndiisdisgfnssvitypdaqlvpgingkaihlvnnessevivhkam dieyndmfnnf
396	AP205 V82C/R80C-hinge-rTT N87	MANKPMQPITSTANKIVWSDPTRLSTTFSASLLRQRVKVGIAELNNVSGQYVSVYKRPAPKP EGCADACVIMPNENQSIctcISGSAENLATLKAEWETHKRNVDTLFASGNAGLGFLDPTAAI VSSDTTgsgEPKSDKTHTPPPAPELLgsgEPKSDKTHTPPPAPELLgsgKnldcwvdneedi dvilkkstilnldinndiisdisgfnssvitypdaqlvpgingkaihlvnnessevivhkam dieyndmfnnf
397	AP205-C65/C69GC-hinge-rTT N87	MANKPMQPITSTANKIVWSDPTRLSTTFSASLLRQRVKVGIAELNNVSGQYVSVYKRPAPKP EGCADAgCVIMPNENQSIRTVISGSAENLATLKAEWETHKRNVDTLFASGNAGLGFLDPTAA IVSSDTTgsgEPKSDKTHTPPPAPELLgsgEPKSDKTHTPPPAPELLgsgKnldcwvdneed idvilkkstilnldinndiisdisgfnssvitypdaqlvpgingkaihlvnnessevivhka mdieyndmfnnf

The fusion protein of the self-assembling protein nanoparticle carrier can include various tags and sequences for production and purification. Typically such protein tags are linked to the N- or C-terminus of the monomer and are ultimately removed (for example by selective protease cleave) from the monomer. For production in cells, the fusion protein of the self-assembling protein nanoparticle carrier can further include a signal peptide that is cleaved off during cellular processing. The fusion proteins can be expressed in appropriate cells (e.g., HEK 293 Freestyle cells) and the fusion proteins are secreted from the cells and self-assemble into the protein nanoparticle carrier. The protein nanoparticle carrier can be purified using known

techniques, for example by a few different chromatography procedures, *e.g.* Mono Q (anion exchange) followed by size exclusion (SUPEROSE® 6) chromatography.

B. HIV-1 Env Fusion Peptide

Any combination of HIV-1 Env fusion peptide and self-assembling protein nanoparticle carrier may be selected from the specific HIV-1 Env fusion peptide and self-assembling protein nanoparticle carriers provided herein to generate the immunogenic conjugate.

HIV-1 can be classified into four groups: the "major" group M, the "outlier" group O, group N, and group P. Within group M, there are several genetically distinct clades (or subtypes) of HIV-1. The HIV-1 Env fusion peptide included in the immunogenic conjugate can be the fusion peptide from any subtype of HIV, such as groups M, N, O, or P or clade A, B, C, D, F, G, H, J or K and the like.

The HIV-1 Env fusion peptide included in the immunogenic conjugate can consist essentially of or consist of residue 512 to one of residues 514-521 (such as residues 512-519) of HIV-1 Env (HXB2) numbering of the Env protein from any subtype of HIV, such as groups M, N, O, or P or clade A, B, C, D, F, G, H, J or K and the like. In some embodiments, the HIV-1 Env fusion peptide included in the immunogenic conjugate can consist essentially of or consist of residue 512 to one of residues 515-521 of HIV-1 Env (HXB2) numbering of the Env protein from any subtype of HIV, such as groups M, N, O, or P or clade A, B, C, D, F, G, H, J or K and the like. In some embodiments, The HIV-1 Env fusion peptide included in the immunogenic conjugate can consist essentially of or consist of residue 512 to one of residues 516-521 of HIV-1 Env (HXB2) numbering of the Env protein from any subtype of HIV, such as groups M, N, O, or P or clade A, B, C, D, F, G, H, J or K and the like. HIV Env fusion peptides from the different HIV clades, as well as nucleic acid sequences encoding such proteins and methods for the manipulation and insertion of such nucleic acid sequences into vectors, are known (see, e.g., HIV Sequence Compendium, Division of AIDS, National Institute of Allergy and Infectious Diseases (2003); HIV Sequence Database (hiv-web.lanl.gov/content/hiv-db/mainpage.html); Sambrook et al., Molecular Cloning, a Laboratory Manual, 2d edition, Cold Spring Harbor Press, Cold Spring Harbor, N. Y. (1989); Ausubel et al., Current Protocols in Molecular Biology, Greene Publishing Associates and John Wiley & Sons, New York, N. Y. (1994)).

In some embodiments, the HIV-1 Env fusion peptides included in the immunogenic conjugate consists essentially of or consists of from 5 to 10 residues (such as 5, 6, 7, 8, 9, or 10 residues or 7-9 residues or 8-10 residues or 6-8 residues) from the N-terminus of the amino acid sequence set forth as AVGIGAVFLG (SEQ ID NO: 1). These residues correspond to HIV-1 Env positions 512-521 (HXB2 numbering). In some embodiments, the HIV-1 Env fusion peptides included in the immunogenic conjugate consists essentially of or consists of the amino acid sequence set forth as residues 1-8 of SEQ ID NO:1.

In some embodiments, the HIV-1 Env fusion peptides included in the immunogenic conjugate consists essentially of or consists of from 5 to 10 residues (such as 5, 6, 7, 8, 9, or 10 residues or 7-9 residues or 8-10 residues or 6-8 residues) from the N-terminus of the amino acid sequence set forth as

AVGLGAVFLG (SEQ ID NO: 2). These residues correspond to HIV-1 Env positions 512-521 (HXB2 numbering). In some embodiments, the HIV-1 Env fusion peptides included in the immunogenic conjugate consists essentially of or consists of the amino acid sequence set forth as residues 1-8 of SEQ ID NO: 2.

In some embodiments, the HIV-1 Env fusion peptides included in the immunogenic conjugate consists essentially of or consists of from 5 to 10 residues (such as 5, 6, 7, 8, 9, or 10 residues or 7-9 residues or 8-10 residues or 6-8 residues) from the N-terminus of the amino acid sequence set forth as AVGIGAMIFG (SEQ ID NO: 3). These residues correspond to HIV-1 Env positions 512-521 (HXB2 numbering). In some embodiments, the HIV-1 Env fusion peptides included in the immunogenic conjugate consists essentially of or consists of the amino acid sequence set forth as residues 1-8 of SEQ ID NO: 3.

In some embodiments, the HIV-1 Env fusion peptides included in the immunogenic conjugate consists essentially of or consists of from 5 to 11 residues (such as 5, 6, 7, 8, 9, 10, or 11 residues or 7-9 residues or 8-10 residues or 6-8 residues) from the N-terminus of the amino acid sequence set forth as AVGTIGAMFLG (SEQ ID NO: 4). These residues correspond to HIV-1 Env positions 512-521 (HXB2 numbering). In some embodiments, the HIV-1 Env fusion peptides included in the immunogenic conjugate consists essentially of or consists of the amino acid sequence set forth as residues 1-8 of SEQ ID NO: 4.

In some embodiments, the HIV-1 Env fusion peptides included in the immunogenic conjugate consists essentially of or consists of from 5 to 10 residues (such as 5, 6, 7, 8, 9, or 10 residues or 7-9 residues or 8-10 residues or 6-8 residues) from the N-terminus of the amino acid sequence set forth as AVGIGAMFLG (SEQ ID NO: 5). These residues correspond to HIV-1 Env positions 512-521 (HXB2 numbering). In some embodiments, the HIV-1 Env fusion peptides included in the immunogenic conjugate consists essentially of or consists of the amino acid sequence set forth as residues 1-8 of SEQ ID NO: 5.

In some embodiments, the HIV-1 Env fusion peptides included in the immunogenic conjugate consists essentially of or consists of from 5 to 10 residues (such as 5, 6, 7, 8, 9, or 10 residues or 7-9 residues or 8-10 residues or 6-8 residues) from the N-terminus of the amino acid sequence set forth as AVGIGALFLG (SEQ ID NO: 6). These residues correspond to HIV-1 Env positions 512-521 (HXB2 numbering). In some embodiments, the HIV-1 Env fusion peptides included in the immunogenic conjugate consists essentially of or consists of the amino acid sequence set forth as residues 1-8 of SEQ ID NO: 6.

In some embodiments, the HIV-1 Env fusion peptides included in the immunogenic conjugate consists essentially of or consists of from 5 to 10 residues (such as 5, 6, 7, 8, 9, or 10 residues or 7-9 residues or 8-10 residues or 6-8 residues) from the N-terminus of the amino acid sequence set forth as AIGLGAMFLG (SEQ ID NO: 7). These residues correspond to HIV-1 Env positions 512-521 (HXB2 numbering). In some embodiments, the HIV-1 Env fusion peptides included in the immunogenic conjugate consists essentially of or consists of the amino acid sequence set forth as residues 1-8 of SEQ ID NO: 7.

In some embodiments, the HIV-1 Env fusion peptides included in the immunogenic conjugate consists essentially of or consists of from 5 to 10 residues (such as 5, 6, 7, 8, 9, or 10 residues or 7-9 residues or 8-10 residues or 6-8 residues) from the N-terminus of the amino acid sequence set forth as AVGLGAVFIG (SEQ ID NO: 8). These residues correspond to HIV-1 Env positions 512-521 (HXB2

numbering). In some embodiments, the HIV-1 Env fusion peptides included in the immunogenic conjugate consists essentially of or consists of the amino acid sequence set forth as residues 1-8 of SEQ ID NO: 8.

In some embodiments, the HIV-1 Env fusion peptides included in the immunogenic conjugate consists essentially of or consists of from 5 to 10 residues (such as 5, 6, 7, 8, 9, or 10 residues or 7-9 residues or 8-10 residues or 6-8 residues) from the N-terminus of the amino acid sequence set forth as AVGIGAVLLG (SEQ ID NO: 9). These residues correspond to HIV-1 Env positions 512-521 (HXB2 numbering). In some embodiments, the HIV-1 Env fusion peptides included in the immunogenic conjugate consists essentially of or consists of the amino acid sequence set forth as residues 1-8 of SEQ ID NO: 9.

In some embodiments, the HIV-1 Env fusion peptides included in the immunogenic conjugate consists essentially of or consists of from 5 to 10 residues (such as 5, 6, 7, 8, 9, or 10 residues or 7-9 residues or 8-10 residues or 6-8 residues) from the N-terminus of the amino acid sequence set forth as AVGIGAVFIG (SEQ ID NO: 10). These residues correspond to HIV-1 Env positions 512-521 (HXB2 numbering). In some embodiments, the HIV-1 Env fusion peptides included in the immunogenic conjugate consists essentially of or consists of the amino acid sequence set forth as residues 1-8 of SEQ ID NO: 10.

In some embodiments, the HIV-1 Env fusion peptides included in the immunogenic conjugate consists essentially of or consists of from 5 to 10 residues (such as 5, 6, 7, 8, 9, or 10 residues or 7-9 residues or 8-10 residues or 6-8 residues) from the N-terminus of the amino acid sequence set forth as AIGLGALFLG (SEQ ID NO: 11). These residues correspond to HIV-1 Env positions 512-521 (HXB2 numbering). In some embodiments, the HIV-1 Env fusion peptides included in the immunogenic conjugate consists essentially of or consists of the amino acid sequence set forth as residues 1-8 of SEQ ID NO: 11.

In some embodiments, the HIV-1 Env fusion peptides included in the immunogenic conjugate consists essentially of or consists of from 5 to 9 residues (such as 5, 6, 7, 8, or 9 residues or 7-9 residues or 8-9 residues or 6-8 residues) from the N-terminus of the amino acid sequence set forth as AALGAVFLG (SEQ ID NO: 12). These residues correspond to HIV-1 Env positions 512-521 (HXB2 numbering). In some embodiments, the HIV-1 Env fusion peptides included in the immunogenic conjugate consists essentially of or consists of the amino acid sequence set forth as residues 1-8 of SEQ ID NO: 12.

In some embodiments, the immunogenic conjugate comprises any of the above HIV-1 Env fusion peptides (such as AVGIGAVF, residues 1-8 of SEQ ID NO: 1) conjugated to a self-assembling protein nanoparticle carrier formed from fusion proteins comprising a tetanus toxoid heavy chain C fragment and a lumazine synthase nanoparticle subunit, wherein the HIV-1 Env fusion peptides are conjugated to the self-assembling protein nanoparticle carrier by linkers between lysine residues on the self-assembling protein nanoparticle carrier and a heterologous cysteine residue fused to a C-terminal residue of the HIV-1 Env fusion peptides.

In some embodiments, the immunogenic conjugate comprises any of the above HIV-1 Env fusion peptides (such as AVGIGAVF, residues 1-8 of SEQ ID NO: 1) conjugated to a self-assembling protein nanoparticle carrier formed from fusion proteins comprising H influenzae protein D (HiD) and a lumazine synthase nanoparticle subunit, wherein the HIV-1 Env fusion peptides are conjugated to the self-assembling

protein nanoparticle carrier by linkers between lysine residues on the self-assembling protein nanoparticle carrier and a heterologous cysteine residue fused to a C-terminal residue of the HIV-1 Env fusion peptides.

In some embodiments, the immunogenic conjugate comprises any of the above HIV-1 Env fusion peptides (such as AVGIGAVF, residues 1-8 of SEQ ID NO: 1) conjugated to a self-assembling protein nanoparticle carrier formed from fusion proteins comprising diphtheria toxoid or a variant thereof (such as CRM197) and a lumazine synthase nanoparticle subunit, wherein the HIV-1 Env fusion peptides are conjugated to the self-assembling protein nanoparticle carrier by linkers between lysine residues on the self-assembling protein nanoparticle carrier and a heterologous cysteine residue fused to a C-terminal residue of the HIV-1 Env fusion peptides.

In some embodiments, the immunogenic conjugate comprises any of the above HIV-1 Env fusion peptides (such as AVGIGAVF, residues 1-8 of SEQ ID NO: 1) conjugated to a self-assembling protein nanoparticle carrier formed from fusion proteins comprising a tetanus toxoid heavy chain C fragment and a ferritin nanoparticle subunit, wherein the HIV-1 Env fusion peptides are conjugated to the self-assembling protein nanoparticle carrier by linkers between lysine residues on the self-assembling protein nanoparticle carrier and a heterologous cysteine residue fused to a C-terminal residue of the HIV-1 Env fusion peptides.

In some embodiments, the immunogenic conjugate comprises any of the above HIV-1 Env fusion peptides (such as AVGIGAVF, residues 1-8 of SEQ ID NO: 1) conjugated to a self-assembling protein nanoparticle carrier formed from fusion proteins comprising H influenzae protein D (HiD) and a ferritin nanoparticle subunit, wherein the HIV-1 Env fusion peptides are conjugated to the self-assembling protein nanoparticle carrier by linkers between lysine residues on the self-assembling protein nanoparticle carrier and a heterologous cysteine residue fused to a C-terminal residue of the HIV-1 Env fusion peptides.

In some embodiments, the immunogenic conjugate comprises any of the above HIV-1 Env fusion peptides (such as AVGIGAVF, residues 1-8 of SEQ ID NO: 1) conjugated to a self-assembling protein nanoparticle carrier formed from fusion proteins comprising diphtheria toxoid or a variant thereof (such as CRM197) and a ferritin nanoparticle subunit, wherein the HIV-1 Env fusion peptides are conjugated to the self-assembling protein nanoparticle carrier by linkers between lysine residues on the self-assembling protein nanoparticle carrier and a heterologous cysteine residue fused to a C-terminal residue of the HIV-1 Env fusion peptides.

In some embodiments, the immunogenic conjugate comprises any of the above HIV-1 Env fusion peptides (such as AVGIGAVF, residues 1-8 of SEQ ID NO: 1) conjugated to a self-assembling protein nanoparticle carrier formed from fusion proteins comprising a tetanus toxoid heavy chain C fragment and a lumazine synthase nanoparticle subunit and further comprising a heterologous T-cell helper epitope (such as AENLWVTVYYGVPVW (SEQ ID NO: 70) or TEKLWVTVYYGVPVW (SEQ ID NO: 71), wherein the HIV-1 Env fusion peptides are conjugated to the self-assembling protein nanoparticle carrier by linkers between lysine residues on the self-assembling protein nanoparticle carrier and a heterologous cysteine residue fused to a C-terminal residue of the HIV-1 Env fusion peptides.

In some embodiments, the immunogenic conjugate comprises any of the above HIV-1 Env fusion peptides (such as AVGIGAVF, residues 1-8 of SEQ ID NO: 1) conjugated to a self-assembling protein nanoparticle carrier formed from fusion proteins comprising H influenzae protein D (HiD) and a lumazine synthase nanoparticle subunit and further comprising a heterologous T-cell helper epitope (such as AENLWVTVYYGVPVW (SEQ ID NO: 70) or TEKLWVTVYYGVPVW (SEQ ID NO: 71), wherein the HIV-1 Env fusion peptides are conjugated to the self-assembling protein nanoparticle carrier by linkers between lysine residues on the self-assembling protein nanoparticle carrier and a heterologous cysteine residue fused to a C-terminal residue of the HIV-1 Env fusion peptides.

In some embodiments, the immunogenic conjugate comprises any of the above HIV-1 Env fusion peptides (such as AVGIGAVF, residues 1-8 of SEQ ID NO: 1) conjugated to a self-assembling protein nanoparticle carrier formed from fusion proteins comprising diphtheria toxoid or a variant thereof (such as CRM197) and a lumazine synthase nanoparticle subunit and further comprising a heterologous T-cell helper epitope (such as AENLWVTVYYGVPVW (SEQ ID NO: 70) or TEKLWVTVYYGVPVW (SEQ ID NO: 71), wherein the HIV-1 Env fusion peptides are conjugated to the self-assembling protein nanoparticle carrier by linkers between lysine residues on the self-assembling protein nanoparticle carrier and a heterologous cysteine residue fused to a C-terminal residue of the HIV-1 Env fusion peptides.

In some embodiments, the immunogenic conjugate comprises any of the above HIV-1 Env fusion peptides (such as AVGIGAVF, residues 1-8 of SEQ ID NO: 1) conjugated to a self-assembling protein nanoparticle carrier formed from fusion proteins comprising a tetanus toxoid heavy chain C fragment and a ferritin nanoparticle subunit and further comprising a heterologous T-cell helper epitope (such as AENLWVTVYYGVPVW (SEQ ID NO: 70) or TEKLWVTVYYGVPVW (SEQ ID NO: 71), wherein the HIV-1 Env fusion peptides are conjugated to the self-assembling protein nanoparticle carrier by linkers between lysine residues on the self-assembling protein nanoparticle carrier and a heterologous cysteine residue fused to a C-terminal residue of the HIV-1 Env fusion peptides.

In some embodiments, the immunogenic conjugate comprises any of the above HIV-1 Env fusion peptides (such as AVGIGAVF, residues 1-8 of SEQ ID NO: 1) conjugated to a self-assembling protein nanoparticle carrier formed from fusion proteins comprising H influenzae protein D (HiD) and a ferritin nanoparticle subunit and further comprising a heterologous T-cell helper epitope (such as AENLWVTVYYGVPVW (SEQ ID NO: 70) or TEKLWVTVYYGVPVW (SEQ ID NO: 71), wherein the HIV-1 Env fusion peptides are conjugated to the self-assembling protein nanoparticle carrier by linkers between lysine residues on the self-assembling protein nanoparticle carrier and a heterologous cysteine residue fused to a C-terminal residue of the HIV-1 Env fusion peptides.

In some embodiments, the immunogenic conjugate comprises any of the above HIV-1 Env fusion peptides (such as AVGIGAVF, residues 1-8 of SEQ ID NO: 1) conjugated to a self-assembling protein nanoparticle carrier formed from fusion proteins comprising diphtheria toxoid or a variant thereof (such as CRM197) and a ferritin nanoparticle subunit and further comprising a heterologous T-cell helper epitope (such as AENLWVTVYYGVPVW (SEQ ID NO: 70) or TEKLWVTVYYGVPVW (SEQ ID NO: 71),

wherein the HIV-1 Env fusion peptides are conjugated to the self-assembling protein nanoparticle carrier by linkers between lysine residues on the self-assembling protein nanoparticle carrier and a heterologous cysteine residue fused to a C-terminal residue of the HIV-1 Env fusion peptides.

Typically, the HIV-1 Env fusion peptides are conjugated to the self-assembling protein nanoparticle carrier by a linker. Suitable linkers include, but are not limited to, straight or branched-chain carbon linkers, heterocyclic carbon linkers or peptide linkers. For an immunogenic conjugate from two or more constituents, each of the constituents will contain the necessary reactive groups. Representative combinations of such groups are amino with carboxyl to form amide linkages or carboxy with hydroxyl to form ester linkages or amino with alkyl halides to form alkylamino linkages or thiols with thiols to form disulfides or thiols with maleimides or alkylhalides to form thioethers. Hydroxyl, carboxyl, amino and other functionalities, where not present may be introduced by known methods. Likewise, a wide variety of linking groups may be employed. In some cases, the linking group can be designed to be either hydrophilic or hydrophobic in order to enhance the desired binding characteristics of the fusion peptide and the carrier. The covalent linkages should be stable relative to the solution conditions under which the conjugate is subjected.

In some embodiments, the linkers may be joined to the constituent amino acids through their side chains (such as through a disulfide linkage to cysteine) or to the alpha carbon, amino, and/or carboxyl groups of the terminal amino acids.

The procedure for attaching a molecule to a polypeptide varies according to the chemical structure of the molecule. Polypeptides typically contain a variety of functional groups; for example, carboxylic acid (COOH), free amine ($-NH_2$) or sulfhydryl (-SH) groups, which are available for reaction with a suitable functional group on a polypeptide. Alternatively, the polypeptide is derivatized to expose or attach additional reactive functional groups. The derivatization may involve attachment of any of a number of linker molecules such as those available from Pierce Chemical Company, Rockford, IL.

In some embodiments, a sulfosuccinimidyl (4-iodoacetyl)aminobenzoate (Sulfo-SIAB) linker is used to link the HIV-1 Env fusion peptides to the self-assembling protein nanoparticle carrier. In some embodiments an m-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) linker is used to link the HIV-1 Env fusion peptides to the self-assembling protein nanoparticle carrier.

The immunogenic conjugate includes a plurality of HIV-1 Env fusion peptides conjugated to the self-assembling protein nanoparticle carrier. In several embodiments, the conjugation of multiple HIV-1 Env fusion peptides to a single self-assembling protein nanoparticle carrier is possible because the carrier has multiple lysine or cysteine side-chains that can serve as sites of attachment. The amount of HIV-1 Env fusion peptide reacted with the amount of self-assembling protein nanoparticle carrier may vary depending upon the specific HIV-1 Env fusion peptide and the self-assembling protein nanoparticle carrier. The resulting number of HIV-1 Env fusion peptides linked to a single self-assembling protein nanoparticle carrier molecule may vary depending upon the specific HIV-1 Env fusion peptides and the self-assembling protein nanoparticle carrier.

Following conjugation of the HIV-1 Env fusion peptide to the protein nanoparticle carrier, the conjugate can be purified by appropriate techniques. One goal of the purification step is to separate the unconjugated HIV-1 Env fusion peptide or protein nanoparticle carrier from the conjugate. One method for purification, involving ultrafiltration in the presence of ammonium sulfate, is described in U.S. Patent No. 6,146,902. Alternatively, the conjugates can be purified away from unconjugated HIV-1 Env fusion peptide or protein nanoparticle carrier by any number of standard techniques including, for example, size exclusion chromatography, density gradient centrifugation, hydrophobic interaction chromatography, or ammonium sulfate fractionation. See, for example, Anderson *et al.*, *J. Immunol.* 137:1181-86, 1986 and Jennings & Lugowski, *J. Immunol.* 127:1011-18, 1981. The compositions and purity of the conjugates can be determined by GLC-MS and MALDI-TOF spectrometry, for example.

In several embodiments, the disclosed immunogenic conjugates can be formulated into immunogenic composition (such as vaccines), for example by the addition of a pharmaceutically acceptable carrier and/or adjuvant.

It is understood that some variations can be made in the amino acid sequence of a protein without affecting the activity of the protein. Such variations include insertion of amino acid residues, deletions of amino acid residues, and substitutions of amino acid residues. These variations in sequence can be naturally occurring variations or they can be engineered through the use of genetic engineering techniques. Examples of such techniques are found in see, *e.g.,* Sambrook *et al.* (Molecular Cloning: A Laboratory Manual, 4th ed, Cold Spring Harbor, New York, 2012) and Ausubel *et al.* (In Current Protocols in Molecular Biology, John Wiley & Sons, New York, through supplement 104, 2013, both of which are incorporated herein by reference in their entirety. Thus, the sequence of the fusion proteins of the disclosed self-assembling protein nanoparticle carrier can include modifications, such as amino acid substitutions, deletions or insertions, glycosylation and/or covalent linkage to unrelated proteins (e.g., a protein tag), as long as the fusion proteins self-assemble to form the self-assembling protein nanoparticle carrier.

III. Nanoparticles linked to carrier by isopeptide bond

Also provided herein are embodiments of a self-assembling nanoparticle-carrier protein where the nanoparticle is linked to carrier proteins by an isopeptide bond between a first tag on the subunits of the nanoparticle and a second tag on the carrier protein. In one example, the first and second tags are based on the Streptococcus pyogenes fibronectin binding protein Fbab-B, such as in the SpyTag/SpyCatcher fusion system. The sequence of Streptococcus pyogenes fibronectin binding protein Fbab-B is provided as follows:

GAMVDTLSGLSSEQGQSGDMTIEEDSATHIKFSKRDIDGKELAGATMELRDSSGKTISTWISDGQVKDFYLMPGKYTFVE
TAAPDGYEVATAITFTVNEQGQVTVNGKATKGDAHIVMVDAYKPTK (SEQ ID NO: 398)

The final 13 residues of Streptococcus pyogenes fibronectin binding protein Fbab-B can be used as a first tag (the spytag) and the remaining residues of Streptococcus pyogenes fibronectin binding protein Fbab-

B are the second tag (spycatcher). When mixed under appropriate conditions the two Streptococcus pyogenes fibronectin binding protein Fbab-B segments bind and form a covalent isopeptide bond.

Any of the nanoparticle subunits disclosed herein can be linked to any of the carrier proteins disclosed herein using the spytag/spycatcher (or other suitable isopeptide bond tag) to generate a nanoparticle-carrier protein to which one or more vaccine antigens (such as an HIV-1 Env fusion peptide as disclosed herein) can be conjugated. In several embodiments, the spytag/spycatcher (or other suitable isopeptide bond tag) is substituted for the peptide linker separating the nanoparticle subunit and carrier protein.

In some embodiments, a lumazine synthase subunit is fused to a spytag and combined with any of the carrier proteins described herein that has been fused to a corresponding spycatcher tag. Non-limiting examples of lumazine synthase subunits fused to a spytag for use in the disclosed embodiments, include:

LS-SpyTag
AHIVMVDAYKPTKgsgsaMQIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVA
AGELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTLEQAIERAGTKHGNKGWEAALS
AIEMANLFKSLR (SEQ ID NO: 399)

LS-SpyTag LODS3 (single Cysteine?)
AHIVMVDAYKPTKgsgsaMQIYEGKLTAEGLRFGIVASRFNHALVDRLVEGcIDAIVRHGGREEDITLVRVPGSWEIPVA
AGELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTLEQAIERAGTKHGNKGWEAALS
AIEMANLFKSLR (SEQ ID NO: 400)

LS-SpyTag LODS5 (intra-protomer)
AHIVMVDAYKPTKgsgsaMQIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVA
AGELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPIcFGVITADTLEQAIERAGTKHGNKGWEAALc
AIEMANLFKSLR (SEQ ID NO: 401)

LS-SpyTag DS2-49
AHIVMVDAYKPTKgsgsaMcIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDIcLVRVPGSWEIPVA
AGELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTLEQAIERAGTKHGNKGWEAALS
AIEMANLFKSLR (SEQ ID NO: 402)

LS-SpyTag DS54-142
AHIVMVDAYKPTKgsgsaMQIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVcGSWEIPVA
AGELARKEnIsAVIAIGVLIRGATPHFDYIASEVSKGLADLSLELRKPITFGVITADTLEQAIERAGTKHGNKGWEAALc
AIEMANLFKSLR (SEQ ID NO: 403)

LS-SpyTag DS95-101
AHIVMVDAYKPTKgsgsaMQIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDAIVRHGGREEDITLVRVPGSWEIPVA
AGELARKEnIsAVIAIGVLIRGATPHFDYIAScVSKGLcDLSLELRKPITFGVITADTLEQAIERAGTKHGNKGWEAALS
AIEMANLFKSLR (SEQ ID NO: 404)

In some embodiments, a ferritin subunit is fused to a spytag and combined with any of the carrier proteins described herein that has been fused to a corresponding spycatcher tag. Non-limiting examples of ferritin subunits fused to a spytag for use in the disclosed embodiments, include:

Ferr 96N SpyTag N-2-THS
AHIVMVDAYKPTKgggsgDPMLSKDIIKLLNEQVNKEMQSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIF
LNENNVPVQLTSISAPEHKFEGLTQIFQKAYEHEQnISESINNIVDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDK
IELIGNENHGLYLADQYVKGIAKSRKS (SEQ ID NO: 405)

Ferr 148S SpyTag N5-THS

AHIVMVDAYKPTKgggsgDIIKLLNEQVNKEMQSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIFLNENNV
PVQLTSISAPEHKFEGLTQIFQKAYEHEQHISESINNIVDHAIKSKDHATFNFLQWYVAEQHEEEVLFKDILDKIELIGN
EsHGLYLADQYVKGIAKSRKS (SEQ ID NO: 406)

In some embodiments, an encapsulin subunit is fused to a spytag and combined with any of the carrier proteins described herein that has been fused to a corresponding spycatcher tag. A non-limiting examples of an encapsulin subunit fused to a spytag for use in the disclosed embodiments, includes:

EN G53C-R94C - spytag
MEFLKRSFAPLTEKQWQEIDNRAREIFKTQLYGRKFVDVEGPYGWEYAAHPLCEVEVLSDENEVVKWGLRKSLPLIELRA
TFTLDLWELDNLECGKPNVDLSSLEETVRKVAEFEDEVIFRGCEKSGVKGLLSFEERKIECGSTPKDLLEAIVRALSIFS
KDGIEGPYTLVINTDRWINFLKEEAGHYPLEKRVEECLRGGKIITTPRIEDALVVSERGGDFKLILGQDLSIGYEDREKD
AVRLFITETFTFQVVNPEALILLKFgggsgAHIVMVDAYKPTK (SEQ ID NO: 410)

The spycatcher tag can be genetically fused to any of carrier proteins provided herein for subsequent isopeptide bond linkage to a nanoparticle subunit fused to a corresponding spytag. In some embodiments, a peptide linker is included between the carrier protein and spycatcher tag or between the nanoparticle subunit and spytag. In one example, the rTT carrier protein fused to the spycatcher tag comprises an amino acid sequence set forth as:

MKNLDCWVDNEEDIDVILKKSTILNLDINNDIISDISGFNSSVITYPDAQLVPGINGKAIHLVNNESSEVIVHKAMDIEY
NDMFNNFTVSFWLRVPKVSASHLEQYGTNEYSIISSMKKHSLSIGSGWSVSLKGNNLIWTLKDSAGEVRQITFRDLPDKF
NAYLANKWVFITITNDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFCKALNPKEIEK
LYTSYLSITFLRDFWGNPLRYDTEYYLIPVASSSKDVQLKNITDYMYLTNAPSYTNGKLNIYYRRLYNGLKFIIKRYTPN
NEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVGYNAPGIPLYKKMEAVKLRDLKTYSVQLKLYDDKNA
SLGLVGTHNGQIGNDPNRDILIASNWYFNHLKDKILGCDWYFVPTDEGWTNDgsgDSATHIKFSKRDEDGKELAGATMEL
RDSSGKTISTWISDGQVKDFYLYPGKYTFVETAAPDGYEVATAITFTVNEQGQVTVNGKATKGDAHI (SEQ ID NO:
407)

In one example, the spytag (e.g., AHIVMVDAYKPTK, SEQ ID NO: 408) is genetically fused to the self-assembling protein nanoparticle subunit, and the nanoparticle with spytag is produced under standard conditions. The spycatcher tag (e.g., DSATHIKFSKRDEDGKELAGATMELRDSSGKTISTWISDGQVKDFYLYPGKYTFVETAAPDGYEVA TAITFTVNEQGQVTVNGKATKGDAHI, SEQ ID NO: 409) is genetically fused to the carrier protein (e.g., rTT), and the carrier protein with spycatcher is produced under standard conditions. The nanoparticle/spytag and carrier/spycatcher are subsequently mixed under conditions sufficient for the spycatcher/spytag to form an isopeptide bond and covalently link the nanoparticle and carrier proteins. The resulting nanoparticle carrier can be used immediately or stored for subsequent conjugation to one or more vaccine antigens of interest.

IV. Self-assembling protein nanoparticles

Additionally provided herein are novel self-assembling protein nanoparticles and subunits thereof. In some embodiments, the self-assembling protein nanoparticle subunit comprises or consists of any one of

the self-assembling protein nanoparticle subunit discussed above in Section II.A.1 for fusion with a heterologous carrier and generation of an immunogenic conjugate.

In some embodiments, the self-assembling protein nanoparticle subunit is a lumazine synthase nanoparticle subunit comprising cysteine substitutions to introduce one or more non-native disulfide bonds to increase stability of the nanoparticle, wherein the cysteine substitutions comprise 121C and 131C substitutions, 121CG and 131C substitutions, 121GC and 131C substitutions, 7C and 40C substitutions, 3C and 50C substitutions, 82C and 131CG substitutions, 5C and 52C substitutions, or 95C and A101C substitutions, or a combination thereof, wherein residue numbering corresponds to a reference lumazine synthase subunit set forth as SEQ ID NO: 25. In some embodiments, the self-assembling protein nanoparticle subunit is a lumazine synthase nanoparticle subunit comprising or consisting of the amino acid sequence set forth as any one of SEQ ID NOs: 306-312, or an amino acid sequence at least 90% (such as at least 95%, at least 98%, or at least 99%) identical thereto.;

In some embodiments, the self-assembling protein nanoparticle subunit is an encapsulin nanoparticle subunit comprising cysteine substitutions to introduce one or more non-native disulfide bonds to increase stability of the nanoparticle, wherein the cysteine substitutions comprise 53C and 94C substitutions, 53C and 96C substitutions, or 146C and 185C substitutions, or a combination thereof, wherein residue numbering corresponds to a reference lumazine synthase subunit set forth as SEQ ID NO: 43. In some embodiments, the self-assembling protein nanoparticle subunit is an encapsulin nanoparticle subunit comprising or consisting of the amino acid sequence set forth as any one of SEQ ID NOs: 313-315, or an amino acid sequence at least 90% (such as at least 95%, at least 98%, or at least 99%) identical thereto.;

In some embodiments, the self-assembling protein nanoparticle subunit is an acinetobacter phage AP205 nanoparticle subunit comprising cysteine substitutions to introduce one or more non-native disulfide bonds to increase stability of the nanoparticle, wherein the cysteine substitutions comprise a T81C substitution, 53C and 100C substitution, or 82C and 80C substitutions, or a combination thereof, wherein residue numbering corresponds to a reference lumazine synthase subunit set forth as SEQ ID NO: 316. In some embodiments, the self-assembling protein nanoparticle subunit is a *Acinetobacter* phage AP205 protein subunit comprising or consisting of the amino acid sequence set forth as any one of SEQ ID NOs: 317-320, or an amino acid sequence at least 90% (such as at least 95%, at least 98%, or at least 99%) identical thereto.; or

In some embodiments, the self-assembling protein nanoparticle subunit is a Hepatitis B capsid protein nanoparticle subunit comprising cysteine substitutions to introduce one or more non-native disulfide bonds to increase stability of the nanoparticle, wherein the cysteine substitutions comprise 25C and 127C substitutions, 14C and 36C substations, 29C and 127C substitutions, 18C and 36C substitutions, or 29C and 127C substitutions, or a combination thereof, wherein residue numbering corresponds to a reference lumazine synthase subunit set forth as SEQ ID NO: 321. In some embodiments, the self-assembling protein nanoparticle subunit is a Hepatitis B capsid protein subunit comprising or consisting of the amino acid

sequence set forth as any one of SEQ ID NOs: 322-326, or an amino acid sequence at least 90% (such as at least 95%, at least 98%, or at least 99%) identical thereto..

In some embodiments, the self-assembling protein nanoparticle subunit is a ferritin nanoparticle subunit comprising or consisting of the amino acid sequence set forth as any one of SEQ ID NOs: 258-305, or an amino acid sequence at least 90% (such as at least 95%, at least 98%, or at least 99%) identical thereto.

In some embodiments, the recombinant self-assembling nanoparticle subunit is fused to a heterologous carrier protein, such as any of the heterologous carrier proteins discussed above in Section II.A.2 for fusion with a self-assembling protein nanoparticle subunit and generation of an immunogenic conjugate. In some embodiments, the recombinant self-assembling nanoparticle subunit is fused to a tetanus toxin heavy chain C fragment, a diphtheria toxin variant CRM197, and an H influenzae protein D, a Keyhole Limpet Hemocyanin (KLH) functional unit, a Meningococcal outer membrane protein complex protein, an Outer-membrane lipoprotein carrier protein, or a Cholera toxin B subunit. Fusion of the heterologous cattier protein to the recombinant self-assembling nanoparticle subunit can be direct (e.g., vis peptide bond between the nanoparticle subunit and the carrier) or indirect via a peptide linker. Any suitable peptide linker may be used, such as the linkers discussed above discussed above in Section II.A.3 for fusion with a self-assembling protein nanoparticle subunit and generation of an immunogenic conjugate.

Also provided are self-assembled protein nanoparticles formed from the nanoparticle subunits. If the nanoparticle subunit is fused to a heterologous carrier protein, then the self-assembled protein nanoparticle will include multiple copies of the heterologous carrier.

In further embodiments, the self-assembled protein nanoparticle is conjugated to a vaccine antigen.

V. Polynucleotides and Expression

Polynucleotides encoding a disclosed fusion protein that forms a self-assembling protein nanoparticle carrier or self-assembling protein nanoparticle are also provided. These polynucleotides include DNA, cDNA and RNA sequences which encode the fusion protein. The genetic code can be used to construct a variety of functionally equivalent nucleic acids, such as nucleic acids which differ in sequence but which encode the same protein sequence, or encode a conjugate or fusion protein including the nucleic acid sequence.

In several embodiments, the nucleic acid molecule encodes a precursor of a disclosed fusion protein and/or nanoparticle subunit, that, when expressed in cells under appropriate conditions, is processed and self-assembles into the protein nanoparticle carrier or protein nanoparticle. For example, the nucleic acid molecule can encode a N-terminal signal sequence for entry into the cellular secretory system that is proteolytically cleaved in the during processing of the fusion protein.

Exemplary nucleic acids can be prepared by cloning techniques. Examples of appropriate cloning and sequencing techniques, and instructions sufficient to direct persons of skill through many cloning exercises are known (see, *e.g.*, Sambrook *et al.* (Molecular Cloning: A Laboratory Manual, 4th ed, Cold

Spring Harbor, New York, 2012) and Ausubel *et al.* (In Current Protocols in Molecular Biology, John Wiley & Sons, New York, through supplement 104, 2013).

Nucleic acids can also be prepared by amplification methods. Amplification methods include polymerase chain reaction (PCR), the ligase chain reaction (LCR), the transcription-based amplification system (TAS), the self-sustained sequence replication system (3SR). A wide variety of cloning methods, host cells, and *in vitro* amplification methodologies are well known to persons of skill.

The polynucleotides encoding a disclosed fusion protein and/or nanoparticle subunit can include a recombinant DNA which is incorporated into a vector into an autonomously replicating plasmid or virus or into the genomic DNA of a prokaryote or eukaryote, or which exists as a separate molecule (such as a cDNA) independent of other sequences. The nucleotides can be ribonucleotides, deoxyribonucleotides, or modified forms of either nucleotide. The term includes single and double forms of DNA.

Polynucleotide sequences encoding a disclosed fusion protein and/or nanoparticle subunit can be operatively linked to expression control sequences. An expression control sequence operatively linked to a coding sequence is ligated such that expression of the coding sequence is achieved under conditions compatible with the expression control sequences. The expression control sequences include, but are not limited to, appropriate promoters, enhancers, transcription terminators, a start codon (*i.e.*, ATG) in front of a protein-encoding gene, splicing signals for introns, maintenance of the correct reading frame of that gene to permit proper translation of mRNA, and stop codons.

DNA sequences encoding the disclosed fusion protein and/or nanoparticle subunit can be expressed *in vitro* by DNA transfer into a suitable host cell. The cell may be prokaryotic or eukaryotic. The term also includes any progeny of the subject host cell. It is understood that all progeny may not be identical to the parental cell since there may be mutations that occur during replication. Methods of stable transfer, meaning that the foreign DNA is continuously maintained in the host, are known in the art.

Hosts can include microbial, yeast, insect and mammalian organisms. Methods of expressing DNA sequences having eukaryotic or viral sequences in prokaryotes are well known in the art. Non-limiting examples of suitable host cells include bacteria, archea, insect, fungi (for example, yeast), plant, and animal cells (for example, mammalian cells, such as human). Exemplary cells of use include *Escherichia coli, Bacillus subtilis, Saccharomyces cerevisiae*, *Salmonella typhimurium,* SF9 cells, C129 cells, 293 cells, *Neurospora*, and immortalized mammalian myeloid and lymphoid cell lines. Techniques for the propagation of mammalian cells in culture are well-known (see, *e.g.,* Helgason and Miller (Eds.), 2012, Basic Cell Culture Protocols (Methods in Molecular Biology), 4th Ed., Humana Press). Examples of commonly used mammalian host cell lines are VERO and HeLa cells, CHO cells, and WI38, BHK, and COS cell lines, although cell lines may be used, such as cells designed to provide higher expression, desirable glycosylation patterns, or other features. In some embodiments, the host cells include HEK293 cells or derivatives thereof, such as GnTI$^{-/-}$ cells (ATCC® No. CRL-3022), or HEK-293F cells.

Transformation of a host cell with recombinant DNA can be carried out by conventional techniques. In some embodiments, if the host is prokaryotic, such as, but not limited to, *E. coli*, competent cells which

are capable of DNA uptake can be prepared from cells harvested after exponential growth phase and subsequently treated by the $CaCl_2$ method. Alternatively, $MgCl_2$ or RbCl can be used. Transformation can also be performed after forming a protoplast of the host cell if desired, or by electroporation.

When the host is a eukaryote, such methods of transfection of DNA as calcium phosphate coprecipitates, conventional mechanical procedures such as microinjection, electroporation, insertion of a plasmid encased in liposomes, or viral vectors can be used. Eukaryotic cells can also be co-transformed with polynucleotide sequences encoding a disclosed antigen, and a second foreign DNA molecule encoding a selectable phenotype, such as the herpes simplex thymidine kinase gene. Another method is to use a eukaryotic viral vector, such as simian virus 40 (SV40) or bovine papilloma virus, to transiently infect or transform eukaryotic cells and express the protein (see for example, Viral Expression Vectors, Springer press, Muzyczka ed., 2011). Appropriate expression systems such as plasmids and vectors of use in producing proteins in cells including higher eukaryotic cells such as the COS, CHO, HeLa and myeloma cell lines can be utilized.

Modifications can be made to a nucleic acid encoding a disclosed fusion protein and/or nanoparticle subunit without diminishing its biological activity. Some modifications can be made to facilitate the cloning or expression of the fusion protein. Non-limiting examples of such modifications include termination codons, a methionine added at the amino terminus to provide an initiation site, additional amino acids placed on either terminus to create conveniently located restriction sites, or additional amino acids (such as poly His) to aid in purification steps.

VI. Immunogenic Compositions

Immunogenic compositions comprising a disclosed immunogenic conjugate and a pharmaceutically acceptable carrier are also provided. Such pharmaceutical compositions can be administered to subjects by a variety of administration modes, for example, intramuscular, subcutaneous, intravenous, intra-arterial, intra-articular, intraperitoneal, or parenteral routes. IActual methods for preparing administrable compositions are described in more detail in such publications as *Remingtons Pharmaceutical Sciences*, 19th Ed., Mack Publishing Company, Easton, Pennsylvania, 1995.

Thus, an immunogenic conjugate described herein can be formulated with pharmaceutically acceptable carriers to help retain biological activity while also promoting increased stability during storage within an acceptable temperature range. Potential carriers include, but are not limited to, physiologically balanced culture medium, phosphate buffer saline solution, water, emulsions (*e.g.*, oil/water or water/oil emulsions), various types of wetting agents, cryoprotective additives or stabilizers such as proteins, peptides or hydrolysates (*e.g.*, albumin, gelatin), sugars (*e.g.*, sucrose, lactose, sorbitol), amino acids (*e.g.*, sodium glutamate), or other protective agents. The resulting aqueous solutions may be packaged for use as is or lyophilized. Lyophilized preparations are combined with a sterile solution prior to administration for either single or multiple dosing.

Formulated compositions, especially liquid formulations, may contain a bacteriostat to prevent or minimize degradation during storage, including but not limited to effective concentrations (usually ≦1% w/v) of benzyl alcohol, phenol, m-cresol, chlorobutanol, methylparaben, and/or propylparaben. A bacteriostat may be contraindicated for some patients; therefore, a lyophilized formulation may be reconstituted in a solution either containing or not containing such a component.

The immunogenic compositions of the disclosure can contain as pharmaceutically acceptable vehicles substances as required to approximate physiological conditions, such as pH adjusting and buffering agents, tonicity adjusting agents, wetting agents and the like, for example, sodium acetate, sodium lactate, sodium chloride, potassium chloride, calcium chloride, sorbitan monolaurate, and triethanolamine oleate.

The immunogenic composition may optionally include an adjuvant to enhance an immune response of the host. Suitable adjuvants are, for example, toll-like receptor agonists, alum, AlPO4, alhydrogel, Lipid-A and derivatives or variants thereof, oil-emulsions, saponins, neutral liposomes, liposomes containing the vaccine and cytokines, non-ionic block copolymers, and chemokines. Non-ionic block polymers containing polyoxyethylene (POE) and polyxylpropylene (POP), such as POE-POP-POE block copolymers, MPL™ (3-O-deacylated monophosphoryl lipid A; Corixa, Hamilton, IN) and IL-12 (Genetics Institute, Cambridge, MA) may also be used as an adjuvant (Newman et al., 1998, *Critical Reviews in Therapeutic Drug Carrier Systems* 15:89-142). These adjuvants have the advantage in that they help to stimulate the immune system in a non-specific way, thus enhancing the immune response to a pharmaceutical product.

In some embodiments, the immunogenic composition can be provided as a sterile composition. The immunogenic composition typically contains an effective amount of a disclosed immunogenic conjugate and can be prepared by conventional techniques. Typically, the amount of immunogenic conjugate in each dose of the immunogenic composition is selected as an amount which elicits or primes an immune response without significant, adverse side effects. In some embodiments, the immunogenic composition can be provided in unit dosage form for use to elicit or prime an immune response in a subject, for example, to prevent HIV-1 infection in the subject. A unit dosage form contains a suitable single preselected dosage for administration to a subject, or suitable marked or measured multiples of two or more preselected unit dosages, and/or a metering mechanism for administering the unit dose or multiples thereof.

VII. Methods of Inducing an Immune Response

The disclosed immunogenic conjugates and compositions including same can be administered to a subject to induce an immune response to HIV-1 to prevent, inhibit, and/or treat an HIV-1 infection. The immune response can be a protective immune response, for example a response that prevents or reduces subsequent infection with HIV-1. Elicitation of the immune response can also be used to treat or inhibit infection and illnesses associated with HIV-1 infection. Thus, the disclosed immunogenic conjugates and compositions including same can be used in methods of preventing, inhibiting, or treating an HIV-1 infection. In several embodiments, an effective amount of an immunogenic conjugate or composition

including same can be administered to a subject in order to generate a neutralizing immune response to HIV-1.

When inhibiting, treating, or preventing HIV-1 infection, the methods can be used either to avoid infection in an HIV-1 seronegative subject (*e.g.*, by inducing an immune response that protects against HIV-1 infection), or to treat existing infection in an HIV-1 seropositive subject. The HIV-1 seropositive subject may or may not carry a diagnosis of AIDS. Hence in some embodiments the methods involve selecting a subject at risk for contracting HIV-1 infection, or a subject at risk of developing AIDS (such as a subject with HIV-1 infection), and administering a disclosed immunogenic conjugate or composition including same to the subject to elicit an immune response to HIV-1 in the subject.

Treatment of HIV-1 by inhibiting HIV-1 replication or infection can include delaying the development of AIDS in a subject. Treatment of HIV-1 can also include reducing signs or symptoms associated with the presence of HIV-1 (for example, by reducing or inhibiting HIV-1 replication). In some examples, treatment using the methods disclosed herein prolongs the time of survival of the subject.

Typical subjects intended for treatment with the therapeutics and methods of the present disclosure include humans, as well as non-human primates and other animals. To identify subjects for prophylaxis or treatment according to the methods of the disclosure, accepted screening methods are employed to determine risk factors associated with a targeted or suspected disease or condition, or to determine the status of an existing disease or condition in a subject. These screening methods include, for example, conventional work-ups to determine environmental, familial, occupational, and other such risk factors that may be associated with the targeted or suspected disease or condition, as well as diagnostic methods, such as various ELISA and other immunoassay methods to detect and/or characterize HIV-1 infection. These and other routine methods allow the clinician to select patients in need of therapy using the methods and pharmaceutical compositions of the disclosure.

The disclosed immunogenic conjugates and compositions including same can be used in coordinate (or prime-boost) immunization protocols or combinatorial formulations. In certain embodiments, novel combinatorial immunogenic compositions and coordinate immunization protocols employ separate immunogenic conjugate or formulations, each directed toward eliciting an anti-HIV-1 immune response, such as an immune response to HIV-1 Env protein. Separate immunogenic conjugates and compositions including same that elicit the anti-HIV-1 immune response can be combined in a polyvalent immunogenic composition administered to a subject in a single immunization step, or they can be administered separately (in monovalent immunogenic compositions) in a coordinate immunization protocol.

In one embodiment, a suitable immunization regimen includes at least two separate inoculations with one or more immunogenic compositions including a disclosed immunogen, with a second inoculation being administered more than about two, about three to eight, or about four, weeks following the first inoculation. A third inoculation can be administered several months after the second inoculation, and in specific embodiments, more than about five months after the first inoculation, more than about six months to about two years after the first inoculation, or about eight months to about one year after the first inoculation.

Periodic inoculations beyond the third are also desirable to enhance the subject's "immune memory." The adequacy of the immunization parameters chosen, *e.g.*, formulation, dose, regimen and the like, can be determined by taking aliquots of serum from the subject and assaying antibody titers during the course of the immunization program. Alternatively, the T cell populations can be monitored by conventional methods. In addition, the clinical condition of the subject can be monitored for the desired effect, *e.g.*, prevention of HIV-1 infection or progression to AIDS, improvement in disease state (*e.g.*, reduction in viral load), or reduction in transmission frequency to an uninfected partner. If such monitoring indicates that immunization is sub-optimal, the subject can be boosted with an additional dose of immunogenic composition, and the immunization parameters can be modified in a fashion expected to potentiate the immune response.

It is contemplated that there can be several boosts, and that each boost can be a different HIV-1 immunogen. It is also contemplated in some examples that the boost may be the same immunogen as another boost, or the prime.

In some embodiments, the prime comprises administration of an immunogenic conjugate as described herein, and the boost (or boosts) comprises administration a recombinant HIV-1 Env ectodomain trimer that is stabilized in a prefusion mature closed conformation, for example, as described in PCT App. No. PCT/US2015/048729 (incorporated by reference herein in its entirety).

The prime and the boost can be administered as a single dose or multiple doses, for example, two doses, three doses, four doses, five doses, six doses or more can be administered to a subject over days, weeks or months. Multiple boosts can also be given, such one to five, or more. Different dosages can be used in a series of sequential inoculations. For example, a relatively large dose in a primary inoculation and then a boost with relatively smaller doses. The immune response against the selected antigenic surface can be generated by one or more inoculations of a subject.

In several embodiments, the immunogenic conjugate can be administered to the subject simultaneously with the administration of an adjuvant. In other embodiments, the immunogenic conjugate can be administered to the subject after the administration of an adjuvant and within a sufficient amount of time to elicit the immune response.

Determination of effective dosages in this context is typically based on animal model studies followed up by human clinical trials and is guided by administration protocols that significantly reduce the occurrence or severity of targeted disease symptoms or conditions in the subject, or that elicit a desired response in the subject (such as a neutralizing immune response). Suitable models in this regard include, for example, murine, rat, porcine, feline, ferret, non-human primate. Alternatively, effective dosages can be determined using *in vitro* models (for example, immunologic and histopathologic assays). Using such models, ordinary calculations and adjustments can be used to determine an appropriate concentration and dose to administer an effective amount of the composition (for example, amounts that are effective to elicit a desired immune response or alleviate one or more symptoms of a targeted disease). In alternative embodiments, an effective amount or effective dose of the immunogenic conjugate may simply inhibit or

enhance one or more selected biological activities correlated with a disease or condition, as set forth herein, for either therapeutic or diagnostic purposes.

Dosage can be varied by the attending clinician to maintain a desired concentration at a target site. Higher or lower concentrations can be selected based on the mode of delivery, for example, trans-epidermal, rectal, oral, pulmonary, or intranasal delivery versus intravenous or subcutaneous delivery. The actual dosage of disclosed immunogenic conjugate will vary according to factors such as the disease indication and particular status of the subject (for example, the subject's age, size, fitness, extent of symptoms, susceptibility factors, and the like), time and route of administration, other drugs or treatments being administered concurrently, as well as the specific pharmacology of the composition for eliciting the desired activity or biological response in the subject. Dosage regimens can be adjusted to provide an optimum prophylactic or therapeutic response.

A non-limiting range for an effective amount of the disclosed immunogenic conjugate within the methods and immunogenic compositions of the disclosure is about 0.0001 mg/kg body weight to about 10 mg/kg body weight, such as about 0.01 mg/kg, about 0.02 mg/kg, about 0.03 mg/kg, about 0.04 mg/kg, about 0.05 mg/kg, about 0.06 mg/kg, about 0.07 mg/kg, about 0.08 mg/kg, about 0.09 mg/kg, about 0.1 mg/kg, about 0.2 mg/kg, about 0.3 mg/kg, about 0.4 mg/kg, about 0.5 mg/kg, about 0.6 mg/kg, about 0.7 mg/kg, about 0.8 mg/kg, about 0.9 mg/kg, about 1 mg/kg, about 1.5 mg/kg, about 2 mg/kg, about 2.5 mg/kg, about 3 mg/kg, about 4 mg/kg, about 5 mg/kg, or about 10 mg/kg, for example, 0.01 mg/kg to about 1 mg/kg body weight, about 0.05 mg/kg to about 5 mg/kg body weight, about 0.2 mg/kg to about 2 mg/kg body weight, or about 1.0 mg/kg to about 10 mg/kg body weight. In some embodiments, the dosage includes a set amount of a disclosed immunogenic conjugate such as from about 1-300 μg, for example, a dosage of about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, or about 300 μg.

The dosage and number of doses will depend on the setting, for example, in an adult or anyone primed by prior HIV-1 infection or immunization, a single dose may be a sufficient booster. In naïve subjects, in some examples, at least two doses would be given, for example, at least three doses. In some embodiments, an annual boost is given, for example, along with an annual influenza vaccination.

For any application, immunization with a disclosed immunogenic conjugate can be combined with anti-retroviral therapy, such as HAART. Antiretroviral drugs are broadly classified by the phase of the retrovirus life-cycle that the drug inhibits. The therapeutic agents can be administered before, during, concurrent to and/or after retroviral therapy. In some embodiments, the therapeutic agents are administered following a course of retroviral therapy. The disclosed therapeutic agents can be administered in conjunction with nucleoside and nucleotide reverse transcriptase inhibitors (nRTI), non-nucleoside reverse transcriptase inhibitors (NNRTI), protease inhibitors, Entry inhibitors (or fusion inhibitors), Maturation inhibitors, or a broad spectrum inhibitors, such as natural antivirals. Exemplary agents include lopinavir, ritonavir, zidovudine, lamivudine, tenofovir, emtricitabine and efavirenz.

HIV-1 infection does not need to be completely eliminated or reduced or prevented for the methods to be effective. For example, elicitation of an immune response to HIV-1 with one or more of the disclosed

immunogenic conjugates (or an immunization protocol involving a disclosed immunogenic conjugate) can reduce or inhibit HIV-1 infection by a desired amount, for example, by at least 10%, at least 20%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 98%, or even at least 100% (elimination or prevention of detectable HIV-1 infected cells), as compared to HIV-1 infection in the absence of the therapeutic agent. In additional examples, HIV-1 replication can be reduced or inhibited by the disclosed methods. HIV-1 replication does not need to be completely eliminated for the method to be effective. For example, the immune response elicited using one or more of the disclosed immunogens can reduce HIV-1 replication by a desired amount, for example, by at least 10%, at least 20%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 98%, or even at least 100% (elimination or prevention of detectable HIV-1 replication), as compared to HIV-1 replication in the absence of the immune response.

To successfully reproduce itself, HIV-1 must convert its RNA genome to DNA, which is then imported into the host cell's nucleus and inserted into the host genome through the action of HIV-1 integrase. Because HIV-1's primary cellular target, CD4+ T-Cells, can function as the memory cells of the immune system, integrated HIV-1 can remain dormant for the duration of these cells' lifetime. Memory T-Cells may survive for many years and possibly for decades. This latent HIV-1 reservoir can be measured by co-culturing CD4+ T-cells from infected patients with CD4+ T-Cells from uninfected donors and measuring HIV-1 protein or RNA (See, *e.g.*, Archin et al., *AIDS,* **22**:1131–1135, 2008). In some embodiments, the provided methods induce an immune response in the subject that reduces or eliminates of the latent reservoir of HIV-1 infected cells in a subject. For example, a reduction of at least 10%, at least 20%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 98%, or even at least 100% (elimination of detectable HIV-1) of the latent reservoir of HIV-1 infected cells in a subject, as compared to the latent reservoir of HIV-1 infected cells in a subject in the absence of immunization with one or more of the provided immunogenic conjugates.

Following immunization of a subject, serum can be collected from the subject at appropriate time points, frozen, and stored for neutralization testing. Methods to assay for neutralization activity and include, but are not limited to, plaque reduction neutralization (PRNT) assays, microneutralization assays, flow cytometry based assays, single-cycle infection assays (*e.g.*, as described in Martin *et al.* (2003) *Nature Biotechnology* 21:71-76), and pseudovirus neutralization assays (*e.g.*, as described in Georgiev et al. (Science, 340, 751-756, 2013), Seaman et al. (J. Virol., 84, 1439-1452, 2005), and Mascola et al. (J. Virol., 79, 10103-10107, 2005), each of which is incorporated by reference herein in its entirety. In some embodiments, the serum neutralization activity can be assayed using a panel of HIV-1 pseudoviruses as described in Georgiev et al., Science, 340, 751-756, 2013 or Seaman et al. J. Virol., 84, 1439-1452, 2005. Briefly, pseudovirus stocks are prepared by co-transfection of 293T cells with an HIV-1 Env-deficient backbone and an expression plasmid encoding the Env gene of interest. The serum to be assayed is diluted in Dulbecco's modified Eagle medium-10% FCS (Gibco) and mixed with pseudovirus. After 30 min, 10,000 TZM-bl cells are added, and the plates are incubated for 48 hours. Assays are developed with a luciferase

assay system (Promega, Madison, WI), and the relative light units (RLU) are read on a luminometer (Perkin-Elmer, Waltham, MA). To account for background, a cutoff of $ID_{50} \geq 40$ can be used as a criterion for the presence of serum neutralization activity against a given pseudovirus.

In some embodiments, administration of an effective amount of one or more of the disclosed the immunogenic conjugates to a subject elicits a neutralizing immune response in the subject, wherein serum from the subject neutralizes, with an $ID_{50} \geq 40$, at least 10% (such as at least 15%, at least 20%, at least 30%, at least 40%, at least 50%, or at least 70%) of pseudoviruses is a panel of pseudoviruses including the HIV-1 Env proteins listed in Table S5 or Table S6 of Georgiev et al. (Science, 340, 751-756, 2013), or Table 1 of Seaman et al. (J. Virol., 84, 1439-1452, 2005).

EXAMPLES

The following examples are provided to illustrate particular features of certain embodiments, but the scope of the claims should not be limited to those features exemplified.

Example 1

Immunogenic conjugate of HIV-1 Env fusion peptides conjugated to a self-assembling protein nanoparticle carrier

This example illustrates immunogenic conjugates including HIV-1 Env fusion peptides conjugated to a self-assembling protein nanoparticle carrier. The immunogenic conjugate provides a multivalent platform with superior binding capability for engaging HIV-1 Env fusion peptide-directed broadly neutralizing antibodies and can be used, for example, to prime an immune response in a subject that targets the HIV-1 Env fusion peptide epitope.

FIG. 1 illustrates the design of certain embodiments of the immunogenic conjugate. As shown in FIG. 1A, the self-assembling protein nanoparticle carrier is a multimer of fusion proteins, each including a self-assembling protein nanoparticle subunit fused to a heterologous carrier protein. In some embodiments, the fusion protein can further include a T-cell helper epitope (FIG. 1B), which is then included in the self-assembling protein nanoparticle carrier. The location of the T-cell-helper epitope can be varied in the fusion protein. As shown in FIGs. 1C-1I, the HIV-1 Env fusion peptides (FP) are conjugated to the self-assembling protein nanoparticle carrier. FIGs. 1G-1I illustrate additional embodiments that further include a targeting moiety that targets the immune system in a subject to enhance the immune response to the HIV-1 Env fusion peptide on the immunogenic conjugate. The HIV-1 Env fusion peptides and the targeting moiety can be conjugated to any suitable aspect of the self-assembling protein nanoparticle carrier. In some instances, sulfosuccinimidyl (4-iodoacetyl)aminobenzoate (Sulfo-SIAB) conjugation chemistry is used to conjugate the HIV-1 Env fusion peptides and/or the targeting moiety to exposed lysine residues of the self-assembling protein nanoparticle carrier.

Example 2

HIV-1 Env fusion peptide immunization using a nanoparticle format

To illustrate the effectiveness of the nanoparticle format for immunization with the HIV-1 Env fusion peptide, the FP8 peptide (AVGIGAVF, residues 1-8 of SEQ ID NO: 1) was conjugated to KLH nanoparticles or to KLH monomeric subunits (FIG. 2). The resulting conjugates were administrated to mice, and immune sera assessed for binding to BG505 HIV-1 Env trimer. As shown in FIG. 2 the nanoparticle based immunogen elicited a much greater immune response to the HIV-1 Env trimer than the subunit-based immunogen.

Example 3

Nanoparticle-carriers for display of vaccine antigens

This example illustrates self-assembling protein nanoparticles fused to heterologous carrier proteins for display of vaccine antigens.

Using structure based design, protein nanoparticle subunits were selected for genetic fusion with heterologous carrier proteins by a variety of peptide linkers.

The nanoparticle subunits were ferritin subunits, lumazine synthase subunits, encapsulin subunits, DNA starvation/stationary phase protection protein subunits, T4 fibritin subunits, Sulfur Oxygenase Reductase subunits, Bacteriophage Q Beta Capsid protein (qbeta) subunits, Dihydrolipoyl transacetylase protein (e2p) subunits, Phosphopantetheine Adenylyltransferase (6ccq) subunits, Glutamate Synthase (1f52) subunits, Calcium/calmodulin dependent protein kinase IIa (CaMKIIa) C-terminal fragment (5U6Y) subunits, HIV capsid oligomerization domain subunits, Hexamer subunits, acinetobacter phage AP205 subunits, and Hepatitis B capsid subunits.

The heterologous carrier proteins were tetanus toxin heavy chain C fragment (rTT), diphtheria toxin variant CRM197 (CRM197), H influenzae protein D, Keyhole Limpet Hemocyanin (KLH) functional unit, Meningococcal outer membrane protein complex protein, Outer-membrane lipoprotein carrier protein, and Cholera toxin B subunit.

The linkers were an IgG hinge, a camel IgG2a hinge, a CD8 hinge, and a glycine serine linker.

Combinations of self-assembling nanoparticle subunit, heterologous carrier, and linker were assessed computationally for formation of a multimerized protein nanoparticle with the heterologous carrier protein fused to each subunit displayed on the exterior surface of the nanoparticle. These design assays led to the identification of the fusion proteins set forth as SEQ ID NOs: 72-219, 246-257, and 331-397, which self-assemble to form nanoparticle carrier proteins.

To illustrate the nanoparticle-forming capacity of the identified fusion proteins, a fusion protein containing a lumazine synthase nanoparticle subunit fused to rTT by a 20 amino acid peptide linker (LS-20-rTT) was assessed for nanoparticle self-assembly (FIG. 3). The fusion protein is depicted in FIG. 3A and the sequence is set forth as SEQ ID NO: 73. A mammalian expression construct encoding LS-20-rTT was expressed in mammalian cells using a standard protocol for generating lumazine synthase nanoparticles (see,

Zhang *et al.* "X-ray structure analysis and crystallographic refinement of lumazine synthase from the hyperthermophile *Aquifex aeolicus* at 1.6 Å resolution: determinants of thermostability revealed from structural comparisons." *J Mol Biol.*, 306(5):1099-114, 2001 and Duan et al., "Glycan Masking Focuses Immune Responses to the HIV-1 CD4-Binding Site and Enhances Elicitation of VRC01-Class Precursor Antibodies," *Immunity*, 49(2):301-311, 2018, each of which is incorporated by reference herein), and the resulting nanoparticles self-assemble in the tissue culture media. The nanoparticles were purified and separated by size-exclusion chromatography (FIG. 3B) and assessed by electron microscopy (FIG. 3C). As shown, the resulting nanoparticles are uniform and stable, and ready for conjugation with vaccine antigen.

Additionally, a fusion protein containing a lumazine synthase nanoparticle subunit fused to rTT by an IgG hinge linker (LS-hinge2-rTT) was assessed for nanoparticle self assembly (FIG. 4). The fusion protein sequence is set forth as SEQ ID NO: 362. A mammalian expression construct encoding LS-hinge2-rTT was expressed in mammalian cells as above, and the resulting nanoparticles were purified and assessed by electron microscopy. Again, the resulting nanoparticles are uniform and stable, and ready for conjugation with vaccine antigen.

Additionally, a fusion protein containing a phosphopantetheine adenylyltransferase nanoparticle subunit was assessed for nanoparticle self assembly (FIG. 5). This fusion protein contained two carrier proteins: the fusion protein contained H influenzae protein D carrier fused to the phosphopantetheine adenylyltransferase nanoparticle subunit fused to rTT carrier (HiD-6CCQ-rTT). The fusion protein sequence is set forth as SEQ ID NO: 179. A mammalian expression construct encoding HiD-6CCQ-rTT was expressed in mammalian cells as above, and the resulting nanoparticles were purified and assessed by electron microscopy. The observed particles were generally consistent in size and shape with the known phosphopantetheine adenylyltransferase crystal structure (PDB 6CCQ). Again, the resulting nanoparticles are uniform and stable, and ready for conjugation with vaccine antigen.

Example 4

Conjugation of HIV-1 Env fusion peptide to nanoparticle carrier

The following provides a non-limiting example of a method of conjugating a HIV-1 Env fusion peptide (FP8, AVGIGAVF, residues 1-8 of SEQ ID NO: 1) to a self-assembling protein nanoparticle carrier (formed from LS-PADRE-Env31-rTT fusion proteins, SEQ ID NO: 76) via a sulfosuccinimidyl (4-iodoacetyl)aminobenzoate (Sulfo-SIAB) linker. The protocol used to link the fusion peptide to carrier can be performed according to standard methods (see, e.g., Hermanson. Bioconjugation Techniques, 3rd ed., Chap. 6, p. 306-308. Academic Press, 2013). Briefly, the conjugation protocol includes:

Expression of the self-assembling protein nanoparticle carrier

An expression construct encoding the LS-PADRE-Env31-rTT fusion protein (SEQ ID NO: 76) including an N-terminal signal peptide is expressed in HEK 293 Freestyle cells. The fusion proteins are secreted from the cells and self-assemble into the protein nanoparticle carrier in the supernatant. The

resulting protein nanoparticle carrier is purified using chromatography procedures, including anion exchange followed by size exclusion chromatography.

Activation of LS-PADRE-Env31-rTT nanoparticle carrier:

1. Prepare 10mM stock of sulfo-SIAB crosslinker.
2. Prepare a 1 mg/mL LS-PADRE-Env31-rTT nanoparticle carrier stock in conjugation buffer (10% glycerol, 50mM Na/KPO_4 buffer, pH 8.5, 1mM EDTA).
3. Add sulfo-SIAB to LS-PADRE-Env31-rTT nanoparticle carrier using a 1:1 molar ratio of crosslinker to total Lys on LS-PADRE-Env31-rTT nanoparticle carrier
4. Let reaction proceed at 25°C (room temperature) for 1 hr.
5. At 4°C, pass through a 10 ml Zebra Spin Desalting Column, 7K MWCO (Thermofisher) to remove low molecular weight compounds.

Conjugation of peptide to activated carrier:

1. Prepare a 12 mM stock of FP8 peptide.
2. Allow activated LS-PADRE-Env31-rTT nanoparticle carrier to warm up to 25°C (room temperature). Gradually add peptide to activated carrier using a 1:1 (w/w) ratio.
3. Spin for 2 min; use supernatant and discard precipitate.
4. Incubate reaction supernatant at 4°C overnight.
5. Use a 10 ml Zebra Spin Desalting Column, 7K MWCO (Thermofisher) to remove low molecular weight compounds.
6. Dialyze conjugate against 1XPBS.
7. Analyze product: degree of conjugation by mass spectrometry and antigenic properties by Octet.

Following purification of the FP8-LS-PADRE-Env31-rTT nanoparticle carrier conjugate, antigenicity can be assessed by binding to fusion peptide specific antibody VRC34.

The conjugation protocol and chemistry illustrated in this example can readily be extended to other fusion peptide sequences and other carrier proteins.

Example 5

Nanoparticle-carriers conjugated to vaccine antigens and related immunization assays

This example illustrates self-assembling protein nanoparticles fused to a heterologous carrier and conjugated to vaccine antigens (HIV-1 Env fusion peptide) and immunization therewith.

For the assays described in this example, the nanoparticle subunit was linked to the heterologous carrier by isopeptide bond using the spytag/spycatcher linkage system. FIG. 6 depicts the construction and purification protocol. rTT carrier was genetically fused to the spycatcher tag, and the lumazine synthase subunit was genetically fused to the spytag. The sequence of the LS-spytag fusion is provided as SEQ ID NO: 399. The sequence of the rTT-spycatcher fusion is provided as SEQ ID NO: 407. The rTT-spyC fusion

protein was produced and purified, and lumazine synthase nanoparticles formed form the LS-spytag fusion were produced purified. The rTT-spyC fusion protein and the lumazine synthase nanoparticles formed form the LS-spytag fusion were mixed, allowing the spytag/spycatcher proteins to spontaneously join by isopeptide bond, resulting in a lumazine synthase nanoparticle linked to rTT via the spycatcher/tag linker. Subsequently, the FP8 fusion peptide was conjugated to the purified nanoparticle-carrier by a PEG linker.

The structure of rTT-spyC, LS-SpyT, LS-Spy-rTT, and LS-Spy-rTT-FP8 were assessed by EM (FIG. 7). Further, the LS-Spy-rTT-FP8 nanoparticle carrier was assessed for the number of conjugated HIV-1 Env fusion peptides using ITC, and this was compared to the corresponding number of HIV-1 Env fusion peptides conjugated to monomeric rTT (FIG. 8). The results show that each FP-rTT monomer entity has six competent VRC34.01 Fab binding sites, whereas each LS-Spy-rTT-FP8v1/PEG2 nanoparticle carrier has 152 – 402 competent VRC34.01 Fab binding sites. The VRC34.01 antibody specifically binds to the HIV-1 Env fusion peptide.

The immunogenicity of the LS-Spy-rTT-FP8v1/PEG2 nanoparticle carrier was assessed in a mouse model. The immunization protocol is shown in FIG. 9. For the first three immunizations (weeks 0, 3, and 6), mice received a 25 μg dose of either FP8v1-rTT monomer (Groups 1 and 2) or LS-Spy-rTT-FP8v1/PEG2 nanoparticle carrier (Groups 3 and 4). For the following three immunizations, mice received a 25 μg dose of either BG505 DS-SOSIP trimer (Groups 1 and 3) or the BG505 DS-SOSIP trimer conjugated to a lumazine synthase nanoparticle (Groups 2 and 4). BG505 DS-SOSIP trimer is a known HIV-1 Env immunogen described in Kwon et al. ("Crystal structure, conformational fixation and entry related interactions of mature ligand-free HIV-1 Env," Nat Struct Biol., 22(7):522-531, 2015, incorporated by reference herein). For these assays BG505 DS-SOSIP trimer was linked to lumazine synthase nanoparticles by standard conjugation chemistry. Blood was drawn at weeks 0, 2, 5, 8, 11, 14, and 17.

Thus, this immunization assay interrogates the ability of the LS-Spy-rTT-FP8v1/PEG2 nanoparticle carrier to generate an immune response in an animal model, and also whether this construct can prime an immune response for subsequent immunization with HIV-1 Env trimer.

As shown in FIG. 10, the LS-Spy-rTT-FP8v1/PEG2 immunogen elicited a far superior immune response to HIV-1 Env fusion peptide compared to monomeric FP-rTT (FIGs. 10A and 10B), and also provided superior priming for subsequent immunization with the BG505 trimer or BG505 trimer on lumazine synthase particle. These results illustrate the effectiveness of the self-assembled protein nanoparticle carrier fusion for use as a immunization tool.

Example 6

Disulfide-stabilized nanoparticle subunits for nanoparticle carriers

This example illustrates self-assembling protein nanoparticles fused to heterologous carrier proteins for display of vaccine antigens that are modified to contain a non-native disulfide bond to increase retention of the nanoparticle format.

Using structure based design, self-assembling protein nanoparticle subunits were mutated to contain one or more cysteine substitutions to introduce a non-native disulfide bond that stabilizes the corresponding nanoparticle formed by the subunits. Stabilization increases resistance to disassembly of the nanoparticle compared to a corresponding native subunit sequence under similar conditions. The mutations were assessed computationally to determine whether they would form a disulfide bond that would stabilize the resulting nanoparticle.

Based on this assessment, ferritin subunits set forth as SEQ ID NOs: 258-305, lumazine synthase subunits set forth as SEQ ID NOs: 306-312, encapsulin subunits set forth as SEQ ID NOs: 313-315, Acinetobacter phage AP205 subunits set forth as SEQ ID NOs: 317-320, and Hepatitis B capsid subunits set forth as SEQ ID NOs: 322-326, were identified, which self-assemble to form nanoparticles containing one or more non-native disulfide bonds that stabilize that nanoparticle relative to nanoparticles formed from unmodified subunits. Specific examples of the disulfide stabilized protein nanoparticles fused to carrier proteins are provided as SEQ ID NOs: 331-354, 369-387, 394-397.

To illustrate the nanoparticle-forming capacity of subunits containing the indicated disulfide bonds, an encapsulin subunit containing G53C-R94C mutations to introduce a stabilizing disulfide bond was fused to a spytag, expressed in cells and the corresponding self-assembled nanoparticles were purified and mixed with rTT-spycatcher (FIGs. 11-13) to form encapsulin-rTT nanoparticle carriers, with the carrier protein linked to the nanoparticle via the spytag/catcher isopeptide bond. The sequence of the encapsulin G53C-R94C spytag fusion is provided as SEQ ID NO: 410, and the sequence of the rTT-spycatcher is provided as SEQ ID NO: 407. The purified nanoparticle-carrier was conjugated to FP8 fusion protein using a SIAB linker. FIG. 13 shows by EM that the resulting HIV-1 Env fusion peptide nanoparticle carrier is uniform and stable.

It will be apparent that the precise details of the methods or compositions described may be varied or modified without departing from the spirit of the described embodiments. We claim all such modifications and variations that fall within the scope and spirit of the claims below.

It is claimed:

1. An immunogenic conjugate, comprising:

a self-assembling protein-nanoparticle carrier comprising a multimer of fusion proteins, wherein each fusion protein comprises a self-assembling protein nanoparticle subunit fused to a heterologous carrier protein, and wherein the fusion proteins self-assemble to form the self-assembling protein-nanoparticle carrier; and

HIV-1 Env fusion peptides conjugated to the self-assembling protein-nanoparticle carrier, wherein the HIV-1 Env fusion peptides comprise, from the N-terminus, the amino acid sequence of residue 512 to one of residues 514-521 of a human immunodeficiency virus type 1 (HIV-1) Envelope (Env) protein according to the HXB2 numbering system; and

wherein the immunogen elicits an immune response to HIV-1 Env.

2. The immunogenic conjugate of claim 1, wherein the heterologous carrier protein is selected from any one of a tetanus toxin heavy chain C fragment, a diphtheria toxin variant CRM197, an H influenzae protein D, a Keyhole Limpet Hemocyanin (KLH) functional unit, a Meningococcal outer membrane protein complex protein, an Outer-membrane lipoprotein carrier protein, or a Cholera toxin B subunit.

3. The immunogenic conjugate of claim 2, wherein the heterologous carrier protein is the tetanus toxin heavy chain C fragment.

4. The immunogenic conjugate of claim 2 or claim 3, wherein:

the heterologous carrier protein is the tetanus toxin heavy chain C fragment and comprises the amino acid sequence set forth as any one of SEQ ID NOs: 44-64, or an amino acid sequence at least 90% identical thereto;

the heterologous carrier protein is the diphtheria toxin variant CRM197 and comprises the amino acid sequence set forth as SEQ ID NO: 65, or an amino acid sequence at least 90% identical thereto;

the heterologous carrier protein is the H influenzae protein D and comprises the amino acid sequence set forth as any one of SEQ ID NOs: 66-67, or an amino acid sequence at least 90% identical thereto;

the heterologous carrier protein is the Meningococcal outer membrane protein complex protein and comprises the amino acid sequence set forth as SEQ ID NO: 222, or an amino acid sequence at least 90% identical thereto;

the heterologous carrier protein is the Outer-membrane lipoprotein carrier protein and comprises the amino acid sequence set forth as any one of SEQ ID NOs: 223, or an amino acid sequence at least 90% identical thereto; or

the heterologous carrier protein is the Cholera toxin B subunit and comprises the amino acid sequence set forth as any one of SEQ ID NOs: 224, or an amino acid sequence at least 90% identical thereto.

5. The immunogenic conjugate of any one of the prior claims, wherein the self-assembling protein nanoparticle subunit is a ferritin subunit, a lumazine synthase subunit, an encapsulin subunit, a DNA starvation/stationary phase protection protein subunit, a T4 fibritin subunit, a Sulfur Oxygenase Reductase subunit, a Bacteriophage Q Beta Capsid protein (qbeta) subunit, a Dihydrolipoyl transacetylase protein (e2p) subunit, a Phosphopantetheine Adenylyltransferase (6ccq) subunit, a Glutamate Synthase (1f52) subunit, a Calcium/calmodulin dependent protein kinase IIa (CaMKIIa) C-terminal fragment (5U6Y) subunit, a HIV capsid oligomerization domain subunit, a Hexamer subunit, an acinetobacter phage AP205 subunit, or a Hepatitis B capsid subunit.

6. The immunogenic conjugate of claim 5, wherein the self-assembling protein nanoparticle subunit comprises the lumazine synthase subunit or the ferritin subunit.

7. The immunogenic conjugate of claim 5, wherein the self-assembling protein nanoparticle subunit comprises:

the ferritin subunit and comprises or consists of the amino acid sequence set forth as any one of SEQ ID NOs: 14-24 or 258-305, or an amino acid sequence at least 90 % identical thereto;

the lumazine synthase subunit and comprises or consists of the amino acid sequence set forth as any one of SEQ ID NOs: 25-28 or 306-312, or an amino acid sequence at least 90 % identical thereto;

the DNA starvation/stationary phase protection protein subunit and comprises or consists of the amino acid sequence set forth as any one of SEQ ID NOs: 29-31, or an amino acid sequence at least 90 % identical thereto;

the encapsulin subunit and comprises or consists of the amino acid sequence set forth as any one of SEQ ID NOs: 43 or 313-315, or an amino acid sequence at least 90 % identical thereto;

the T4 fibritin subunit and comprises or consists of the amino acid sequence set forth as SEQ ID NO: 42, or an amino acid sequence at least 90 % identical thereto;

the qbeta subunit and comprises or consists of the amino acid sequence set forth as SEQ ID NO: 32, or an amino acid sequence at least 90 % identical thereto;

the e2p subunit and comprises or consists of the amino acid sequence set forth as SEQ ID NO: 33, or an amino acid sequence at least 90 % identical thereto;

the phosphopantetheine adenylyltransferase subunit and comprises or consists of the amino acid sequence set forth as SEQ ID NO: 34, or an amino acid sequence at least 90 % identical thereto;

the glutamate Synthase subunit and comprises or consists of the amino acid sequence set forth as SEQ ID NOs: 35, or an amino acid sequence at least 90 % identical thereto;

the CaMKIIa C-terminal fragment subunit and comprises or consists of the amino acid sequence set forth as any one of SEQ ID NOs: 36-38, or an amino acid sequence at least 90 % identical thereto;

the HIV capsid oligermerization domain subunit and comprises or consists of the amino acid sequence set forth as any one of SEQ ID NOs: 39-40, or an amino acid sequence at least 90 % identical thereto;

the Hexamer subunit and comprises or consists of the amino acid sequence set forth as SEQ ID NO: 41, or an amino acid sequence at least 90 % identical thereto;

the acinetobacter phage AP205 subunit and comprises or consists of the amino acid sequence set forth as any one of SEQ ID NOs: 316-320, or an amino acid sequence at least 90 % identical thereto; or

the Hepatitis B capsid subunit and comprises or consists of the amino acid sequence set forth as any one of SEQ ID NOs: 321-326, or an amino acid sequence at least 90 % identical thereto;.

8. The immunogenic conjugate of any one of the prior claims, wherein the heterologous carrier protein is fused to the subunit of the self-assembling protein nanoparticle by a peptide linker.

9. The immunogenic conjugate of claim 8, wherein the peptide linker is an IgG hinge, a camel IgG2a hinge, a CD8 hinge, or a glycine serine linker.

10. The immunogenic conjugate of claim 8 or claim 9, wherein the peptide linker comprises or consists of an amino acid sequence set forth as any one of SEQ ID NOs: 327-330.

11. The immunogenic conjugate of any one of the prior claims, wherein the fusion protein further comprises at least one heterologous T-cell helper epitope sequence.

12. The immunogenic conjugate of claim 11, wherein the at least one heterologous T-cell helper epitope sequence is selected from any one of SEQ ID NOs: 68-69, 221, or 225-245, or the sequence of HIV-1 Env residues 31-45 according to the HXB2 numbering system (Env31-45 epitope), or a combination of two or more thereof.

13. The immunogenic conjugate of claim 12, wherein the HIV-1 Env residues 31-45 comprise the amino acid sequence set forth as any one of:

AENLWVTVYYGVPVW (SEQ ID NO: 70)

TEKLWVTVYYGVPVW (SEQ ID NO: 71)

14. The immunogenic conjugate of any one of the prior claims, wherein the fusion protein comprises or consists essentially of an amino acid sequence at least 90% identical to any one of SEQ ID NOs: 72-219, 246-257, or 331-397.

15. The immunogenic conjugate of any one of the prior claims, wherein the fusion protein comprises or consists essentially of the amino acid sequence of any one of SEQ ID NOs: 72-219, 246-257, or 331-397.

16. The immunogenic conjugate of any one of the prior claims, wherein the HIV-1 Env fusion peptides comprise, consist essentially of, or consist of HIV-1 Env residues 512-519 (HXB2 numbering).

17. The immunogenic conjugate of any one of the prior claims, wherein the HIV-1 Env fusion peptides comprise, consist essentially of, or consist of the amino acid sequence set forth as any one of:

residues 1-4, 1-5, 1-6, 1-7, 1-8, 1-9, or 1-10 of any one of AVGIGAVFLG (SEQ ID NO: 1), AVGLGAVFLG (SEQ ID NO: 2), AVGIGAMIFG (SEQ ID NO: 3), AVGIGAMFLG (SEQ ID NO: 5), AVGIGALFLG (SEQ ID NO: 6), AIGLGAMFLG (SEQ ID NO: 7), AVGLGAVFIG (SEQ ID NO: 8), AVGIGAVLLG (SEQ ID NO: 9), AVGIGAVFIG (SEQ ID NO: 10), AIGLGALFLG (SEQ ID NO: 11);

residues 1-4, 1-5, 1-6, 1-7, 1-8, 1-9, 1-10, or 1-11 of AVGTIGAMFLG (SEQ ID NO: 4); or

residues 1-4, 1-5, 1-6, 1-7, 1-8, or 1-9 of AALGAVFLG (SEQ ID NO: 12).

18. The immunogenic conjugate of any one of the prior claims, wherein the HIV-1 Env fusion peptide comprises, consists essentially of, or consists of the amino acid sequence set forth as AVGIGAVF (residues 1-8 of SEQ ID NO: 1).

19. The immunogenic conjugate of any one of the prior claims, wherein the HIV-1 Env fusion peptides further comprise a C-terminal cysteine residue.

20. The immunogenic conjugate of any one of the prior claims, wherein the HIV-1 Env fusion peptides are no more than 15 amino acids in length.

21. The immunogenic conjugate of any one of the prior claims, wherein the HIV-1 Env fusion peptides are conjugated to the self-assembling protein nanoparticle carrier by a sulfosuccinimidyl (4-iodoacetyl)aminobenzoate (Sulfo-SIAB) or a m-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) crosslinker.

22. The immunogenic conjugate of any one of the prior claims, further comprising a heterologous moiety that targets the immune system.

23. The immunogenic conjugate of claim 22, wherein the heterologous moiety is a flagellin protein fused to the fusion protein of the self-assembling protein nanoparticle carrier or a pattern recognition receptor agonist conjugated to the self-assembling protein nanoparticle carrier.

24. The immunogenic conjugate of any one of the prior claims, wherein

the HIV-1 Env fusion peptides comprise, consist essentially of, or consist of the amino acid sequence set forth as AVGIGAVF (residues 1-8 of SEQ ID NO: 1); and

the fusion proteins of the self-assembling protein nanoparticle carrier comprise the recombinant tetanus toxin heavy chain C fragment fused to the lumazine synthase subunit and further comprise the heterologous T-cell helper epitope sequence of HIV-1 Env residues 31-45 according to the HXB2 numbering system.

25. A method of producing an immunogenic conjugate, comprising:

expressing a nucleic acid molecule encoding the fusion protein comprising the heterologous carrier protein and the self-assembling protein nanoparticle subunit of any one of claims 1-21 in a host cell;

purifying expressed fusion proteins;

incubating the expressed fusion proteins under conditions sufficient for self-assembly of protein nanoparticle subunits into protein nanoparticles; and

conjugating HIV-1 Env fusion peptides to the protein nanoparticles, wherein the HIV-1 Env fusion peptides comprise, from the N-terminus, the amino acid sequence of residue 512 to one of residues 517-521 of a human immunodeficiency virus type 1 (HIV-1) Envelope (Env) protein according to the HXB2 numbering system; and

purifying the protein nanoparticles conjugated to the HIV-1 Env fusion peptides; thereby producing the immunogenic conjugate.

26. The immunogenic conjugate produced by the method of claim 22.

27. An immunogenic composition comprising the immunogenic conjugate of any of claims 1-24 or 26.

28. A method for generating an immune response to HIV-1 Env in a subject, comprising administering to the subject an effective amount of the immunogenic conjugate or immunogenic composition of any one of claims 1-24 or 26-27 to generate the immune response.

29. The method of claim 28, wherein the immune response treats or inhibits HIV-1 infection in the subject.

30. The method of claim 28-29, wherein administration of the immunogenic conjugate or immunogenic composition to the subject primes a protective immune response to HIV-1 infection in the subject.

31. Use of the immunogenic conjugate or immunogenic composition of any one of claims 1-24 or 26-27 to elicit an immune response to HIV-1 Env in a subject.

32. A recombinant self-assembling nanoparticle subunit, comprising:

a lumazine synthase nanoparticle subunit comprising cysteine substitutions to introduce one or more non-native disulfide bonds to increase stability of the nanoparticle, wherein the cysteine substitutions comprise 121C and 131C substitutions, 121CG and 131C substitutions, 121GC and 131C substitutions, 7C and 40C substitutions, 3C and 50C substitutions, 82C and 131CG substitutions, 5C and 52C substitutions, or 95C and A101C substitutions, or a combination thereof, wherein residue numbering corresponds to a reference lumazine synthase subunit set forth as SEQ ID NO: 25;

an encapsulin nanoparticle subunit comprising cysteine substitutions to introduce one or more non-native disulfide bonds to increase stability of the nanoparticle, wherein the cysteine substitutions comprise 53C and 94C substitutions, 53C and 96C substitutions, or 146C and 185C substitutions, or a combination thereof, wherein residue numbering corresponds to a reference lumazine synthase subunit set forth as SEQ ID NO: 43;

an acinetobacter phage AP205 nanoparticle subunit comprising cysteine substitutions to introduce one or more non-native disulfide bonds to increase stability of the nanoparticle, wherein the cysteine substitutions comprise a T81C substitution, 53C and 100C substitution, or 82C and 80C substitutions, or a combination thereof, wherein residue numbering corresponds to a reference lumazine synthase subunit set forth as SEQ ID NO: 316; or

a Hepatitis B capsid protein nanoparticle subunit comprising cysteine substitutions to introduce one or more non-native disulfide bonds to increase stability of the nanoparticle, wherein the cysteine substitutions comprise 25C and 127C substitutions, 14C and 36C substations, 29C and 127C substitutions, 18C and 36C substitutions, or 29C and 127C substitutions, or a combination thereof, wherein residue numbering corresponds to a reference lumazine synthase subunit set forth as SEQ ID NO: 321.

33. A recombinant self-assembling nanoparticle subunit, comprising:

a ferritin nanoparticle subunit comprising or consisting of the amino acid sequence set forth as any one of SEQ ID NOs: 258-305;

a lumazine synthase nanoparticle subunit comprising or consisting of the amino acid sequence set forth as any one of SEQ ID NOs: 306-312;

an encapsulin nanoparticle subunit comprising or consisting of the amino acid sequence set forth as any one of SEQ ID NOs: 313-315;

a *Acinetobacter* phage AP205 protein subunit comprising or consisting of the amino acid sequence set forth as any one of SEQ ID NOs: 317-320; or

a Hepatitis B capsid protein subunit comprising or consisting of the amino acid sequence set forth as any one of SEQ ID NOs: 322-326.

34. The immunogenic conjugate of claim 33, wherein the recombinant self-assembling nanoparticle subunit is fused to a heterologous carrier protein.

35. The immunogenic conjugate of claim 34, wherein the heterologous carrier protein is selected from any one of a tetanus toxin heavy chain C fragment, a diphtheria toxin variant CRM197, and an H influenzae protein D, a Keyhole Limpet Hemocyanin (KLH) functional unit, a Meningococcal outer membrane protein complex protein, an Outer-membrane lipoprotein carrier protein, or a Cholera toxin B subunit.

36. The immunogenic conjugate of claim 35, wherein the heterologous carrier protein is the tetanus toxin heavy chain C fragment.

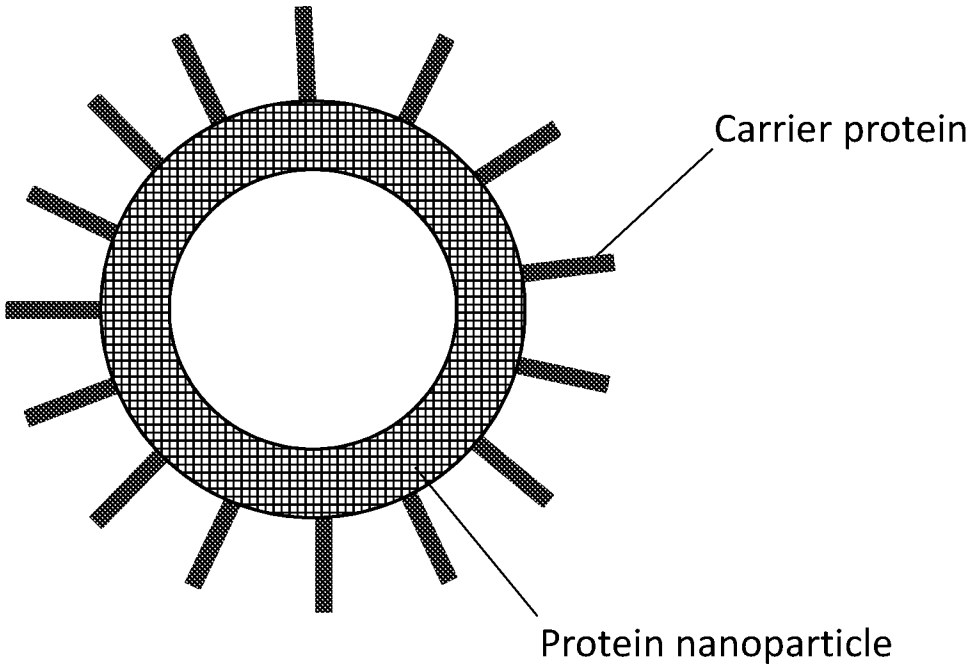
37. A nucleic acid molecule encoding the recombinant self-assembling nanoparticle subunit of any one of claims 32-36.

38. A recombinant self-assembling nanoparticle comprising the recombinant self-assembling nanoparticle subunit of any one of claims 32-37.

39. The recombinant self-assembling nanoparticle of claim 38, conjugated to a vaccine antigen.

FIG. 1A

Self-Assembling Protein Nanoparticle Carrier



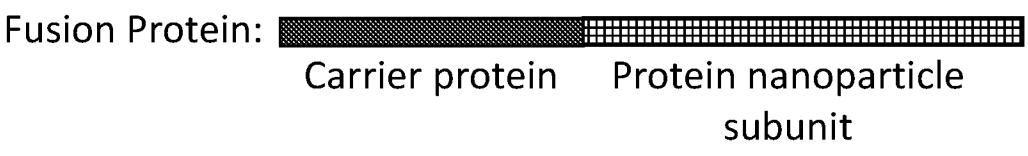
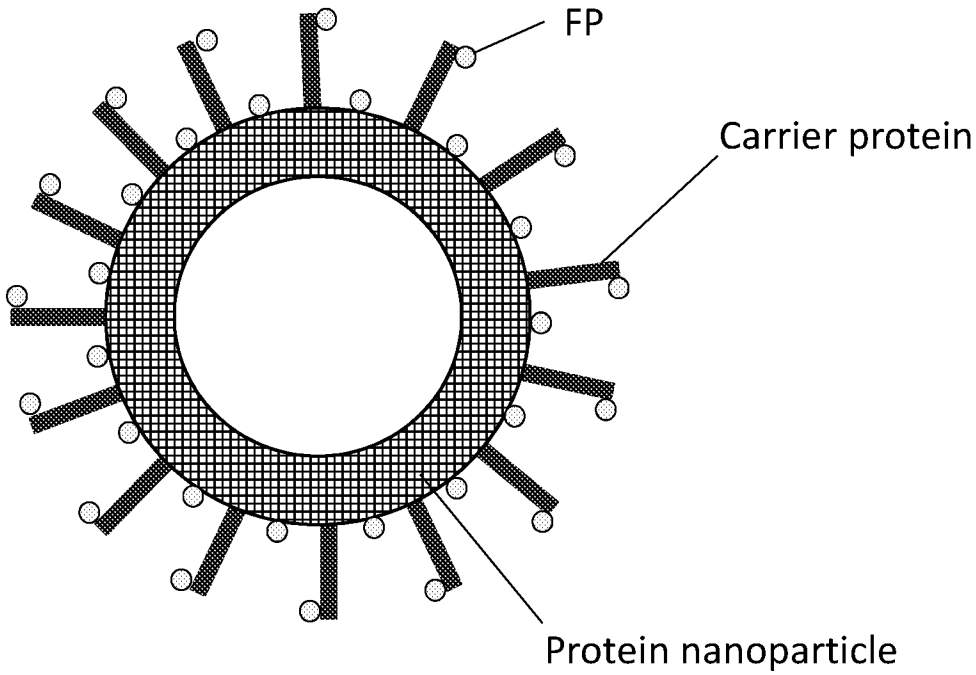


FIG. 1B

Self-Assembling Protein Nanoparticle Carrier
conjugated to HIV-1 Env fusion peptides



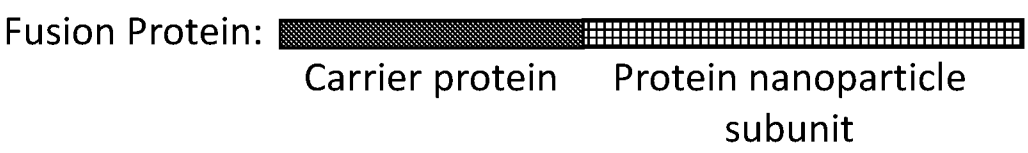
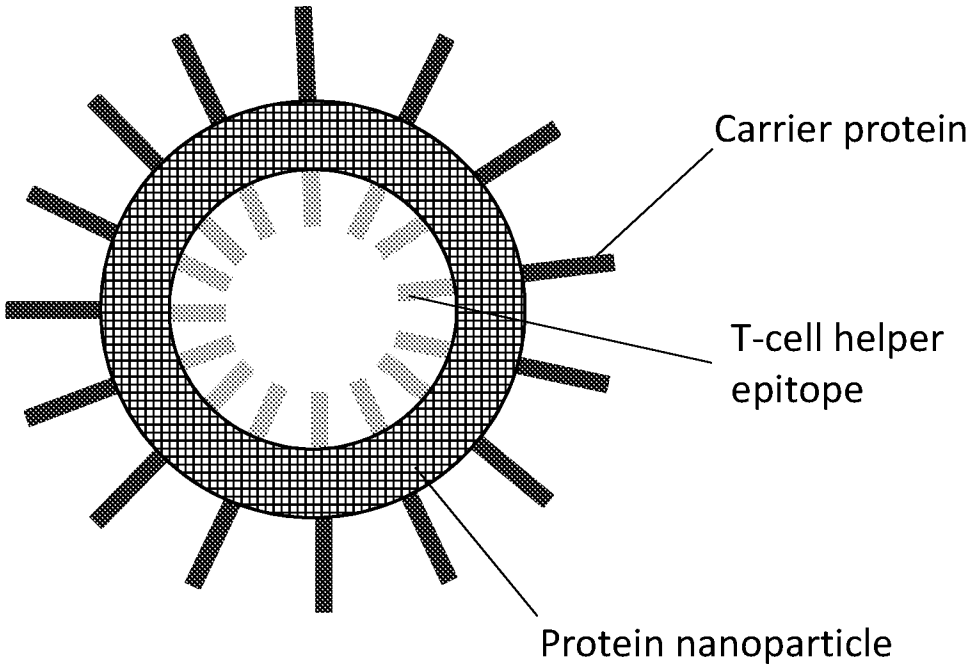


FIG. 1C

Self-Assembling Protein Nanoparticle Carrier
with T-Cell Helper epitopes



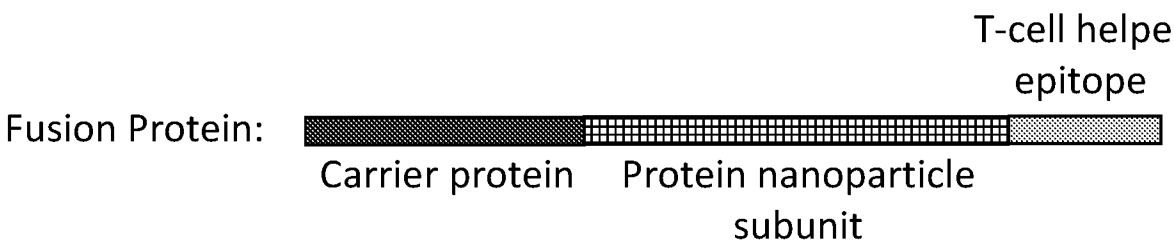
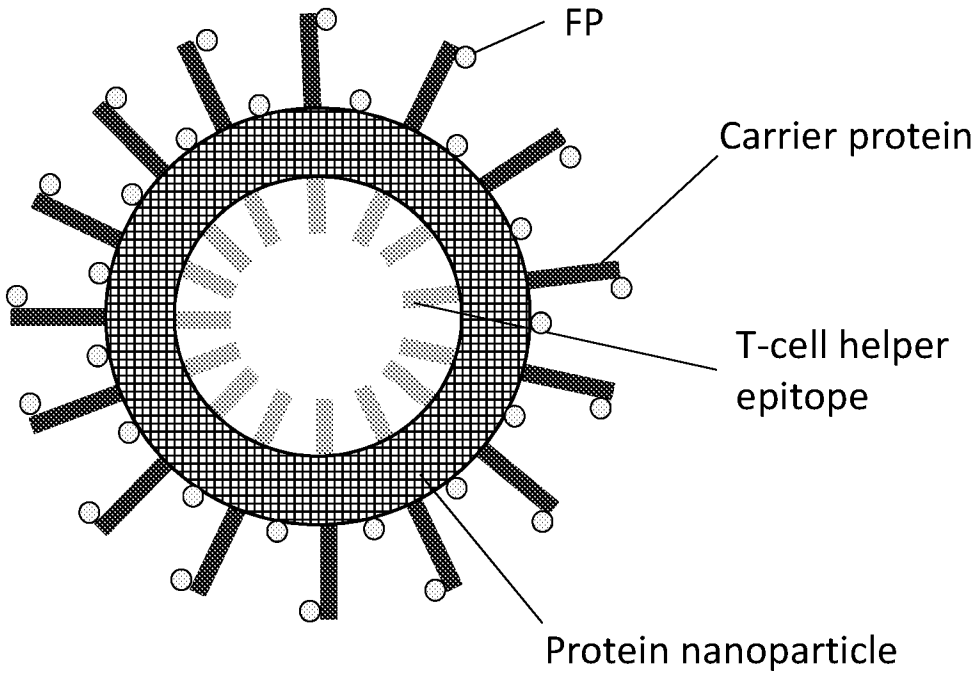


FIG. 1D

Self-Assembling Protein Nanoparticle Carrier
with T-Cell Helper epitopes
conjugated to HIV-1 Env fusion peptides



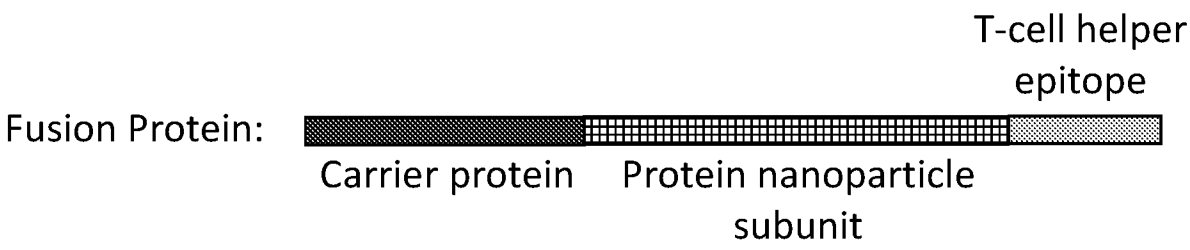
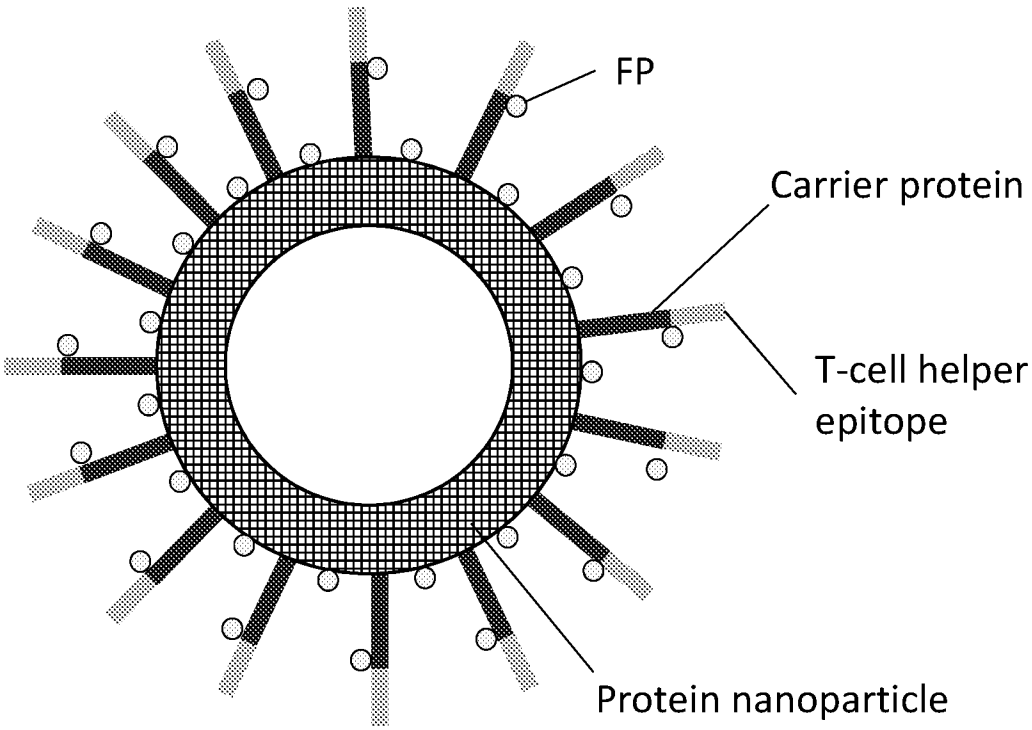


FIG. 1E

Self-Assembling Protein Nanoparticle Carrier
with T-Cell Helper epitopes
conjugated to HIV-1 Env fusion peptides



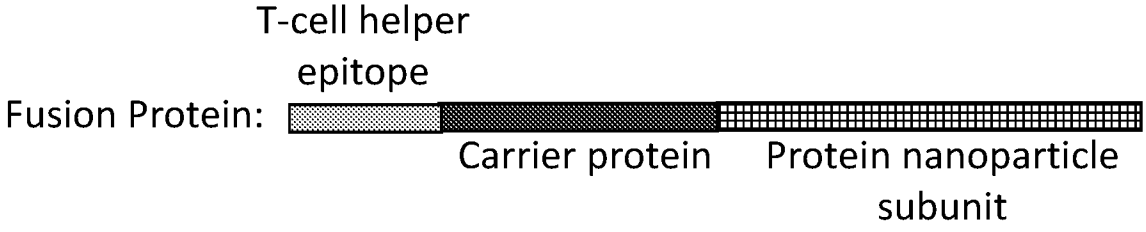
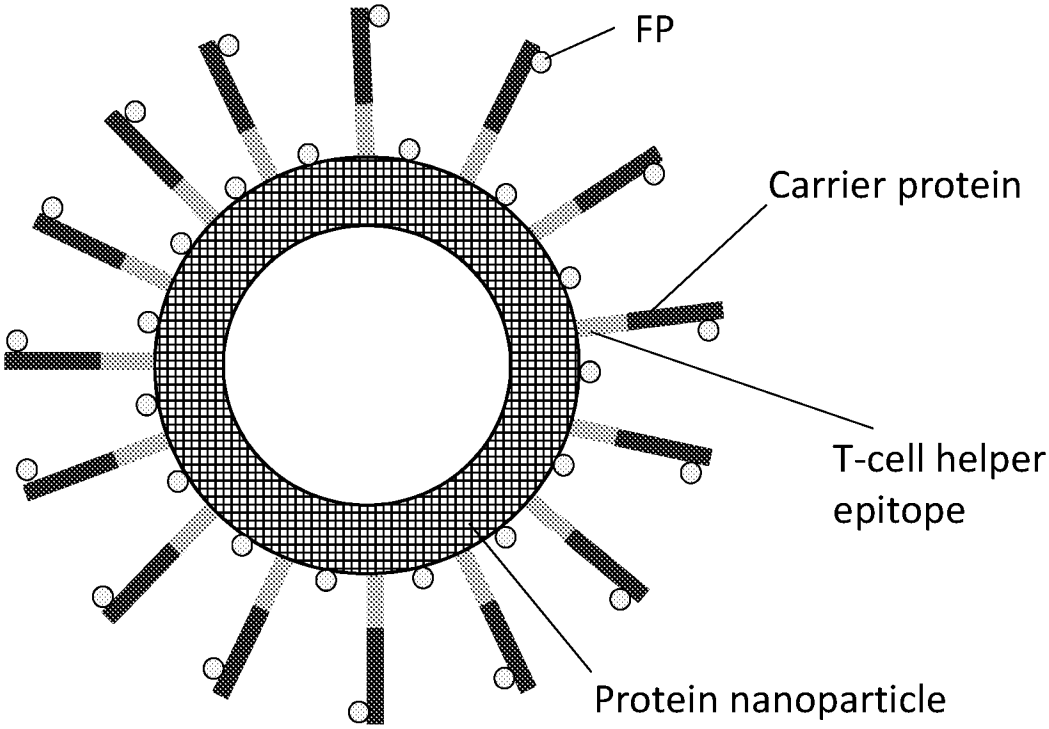


FIG. 1F

Self-Assembling Protein Nanoparticle Carrier
with T-Cell Helper epitopes
conjugated to HIV-1 Env fusion peptides



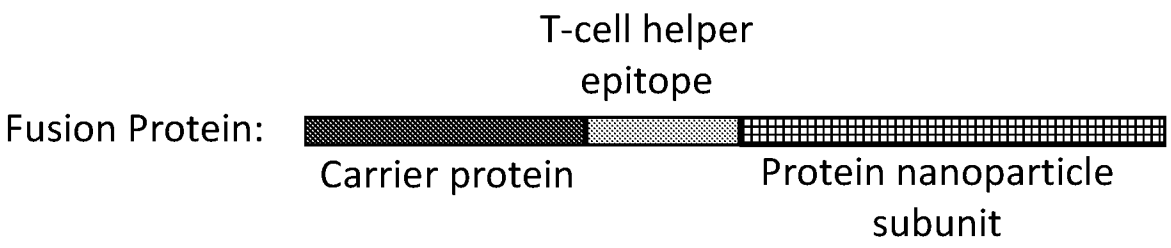
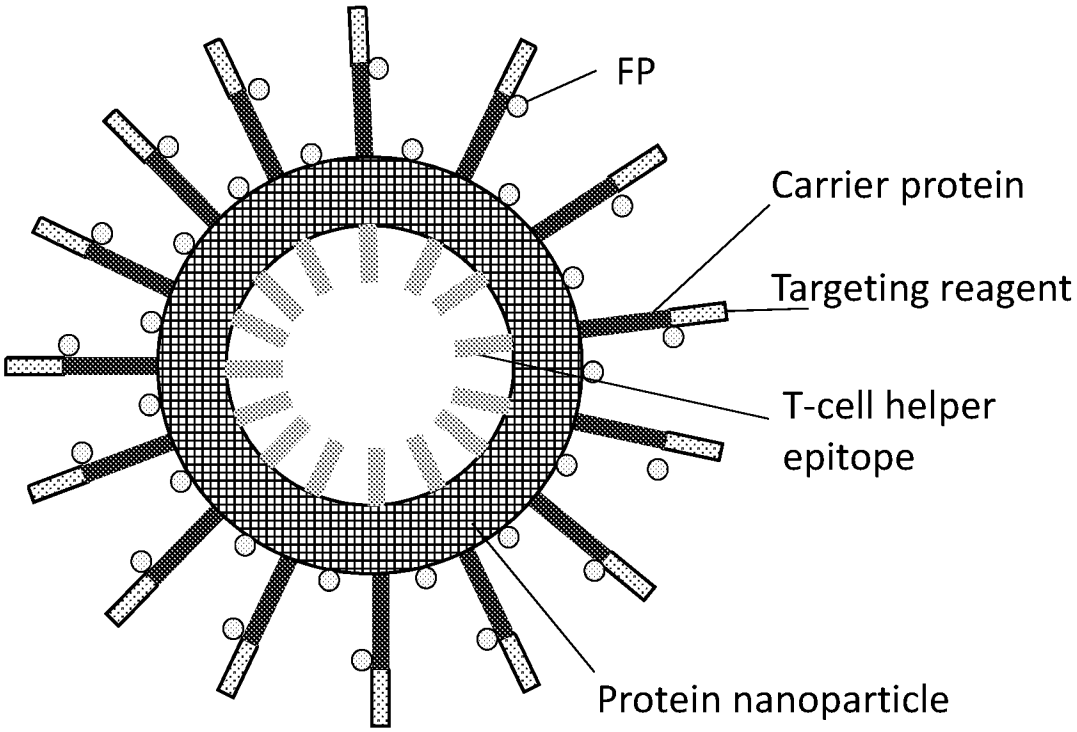


FIG. 1G

Self-Assembling Protein Nanoparticle Carrier
with targeting reagent and T-cell helper
epitope
conjugated to HIV-1 Env fusion peptides



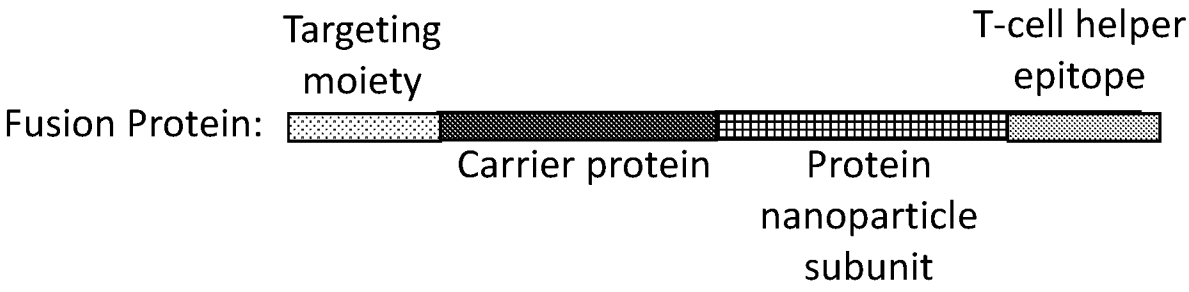
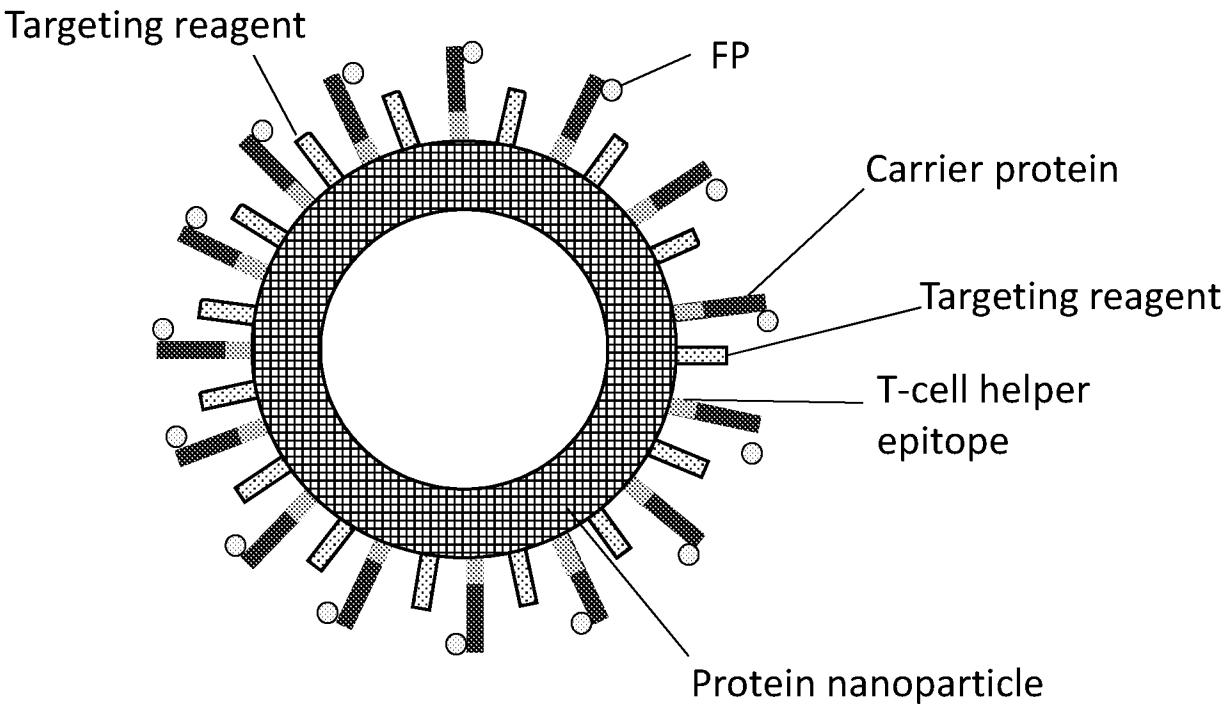


FIG. 1H

Self-Assembling Protein Nanoparticle Carrier
with targeting reagent and T-cell helper epitope
conjugated to HIV-1 Env fusion peptides



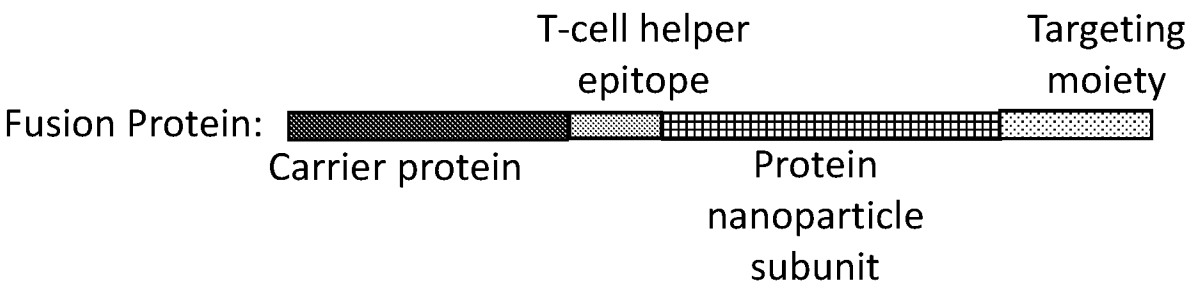
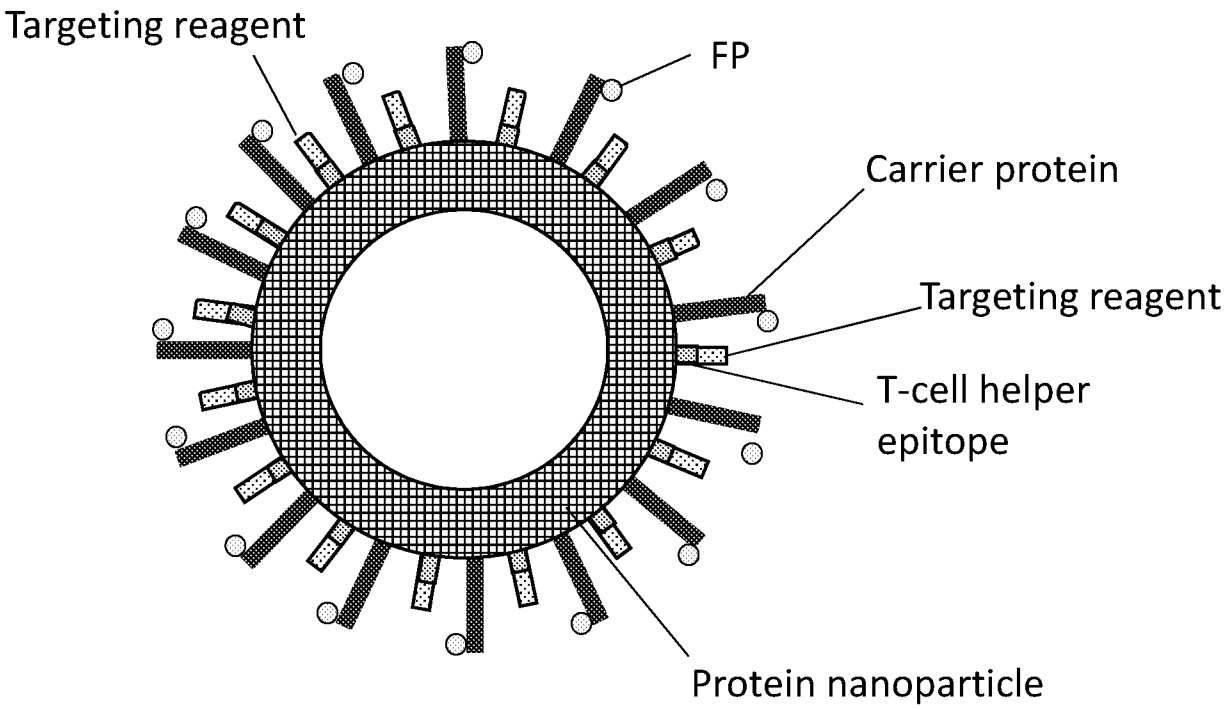


FIG. 1I

Self-Assembling Protein Nanoparticle Carrier
with targeting reagent and T-cell helper epitope
conjugated to HIV-1 Env fusion peptides



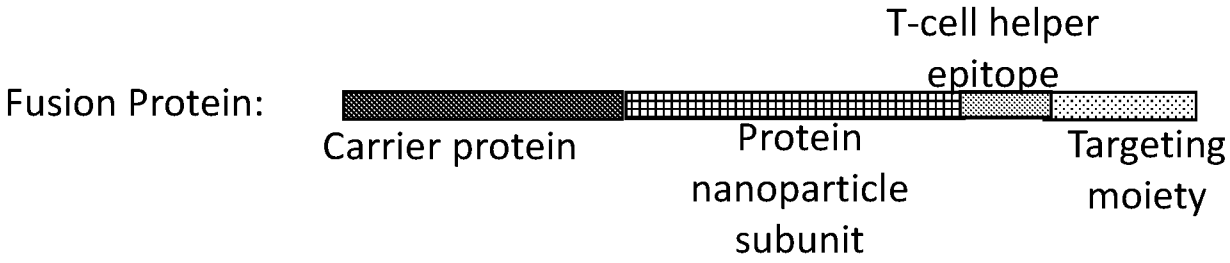
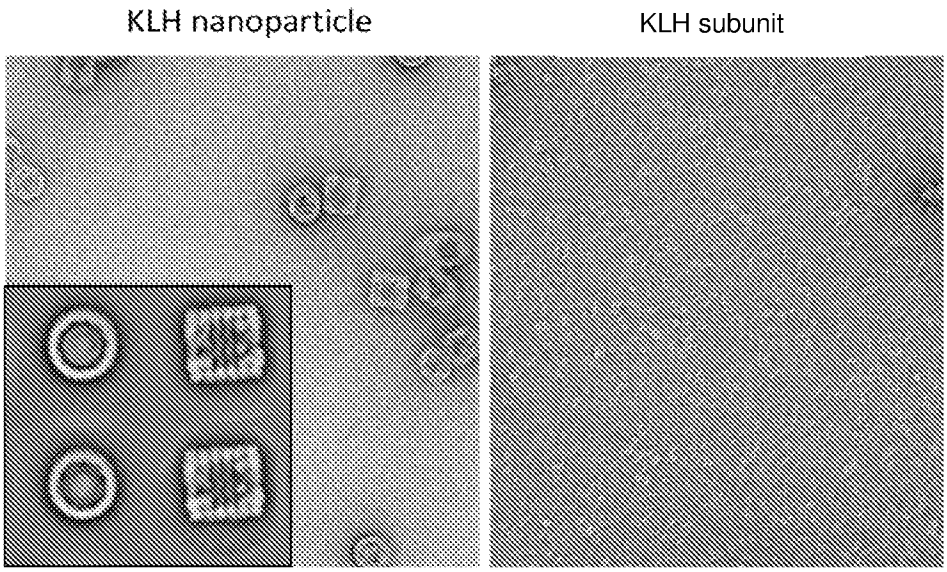


FIG. 2


KLH nanoparticle
KLH subunit

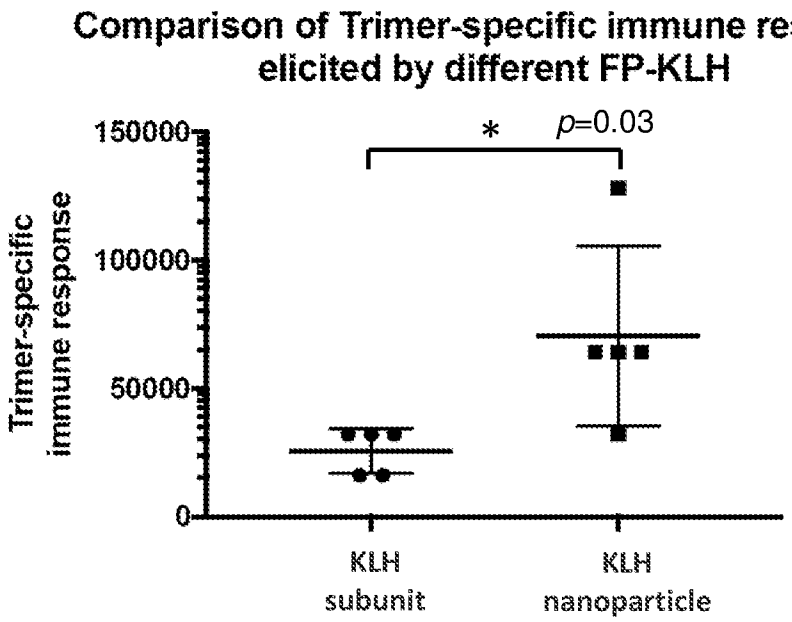

Comparison of Trimer-specific immune response elicited by different FP-KLH
*
p=0.03
150000
100000
50000
0
Trimer-specific immune response
KLH subunit
KLH nanoparticle

FIG. 3A

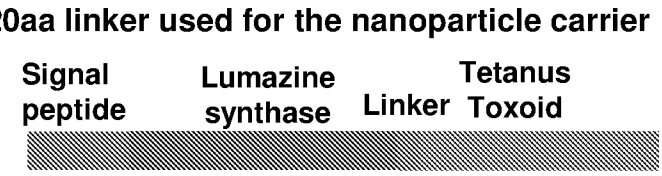


FIG. 3B

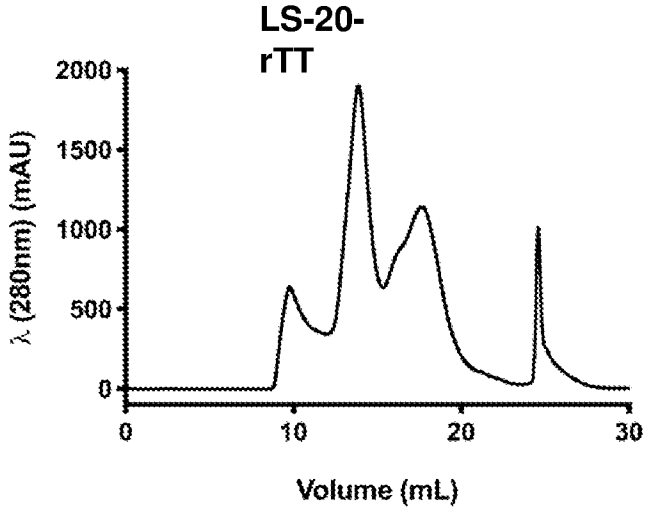


FIG. 3C

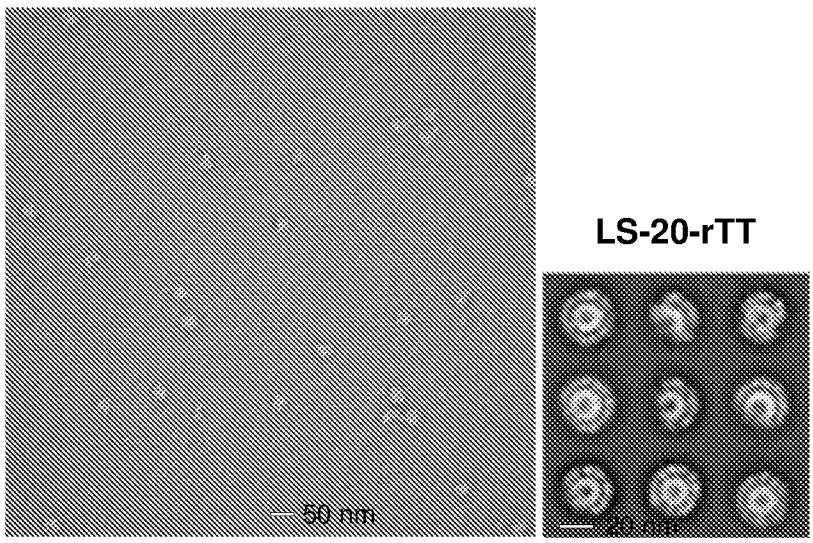
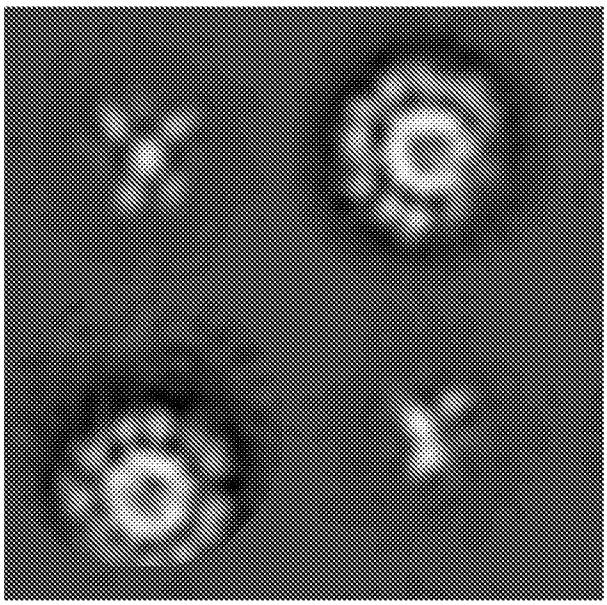


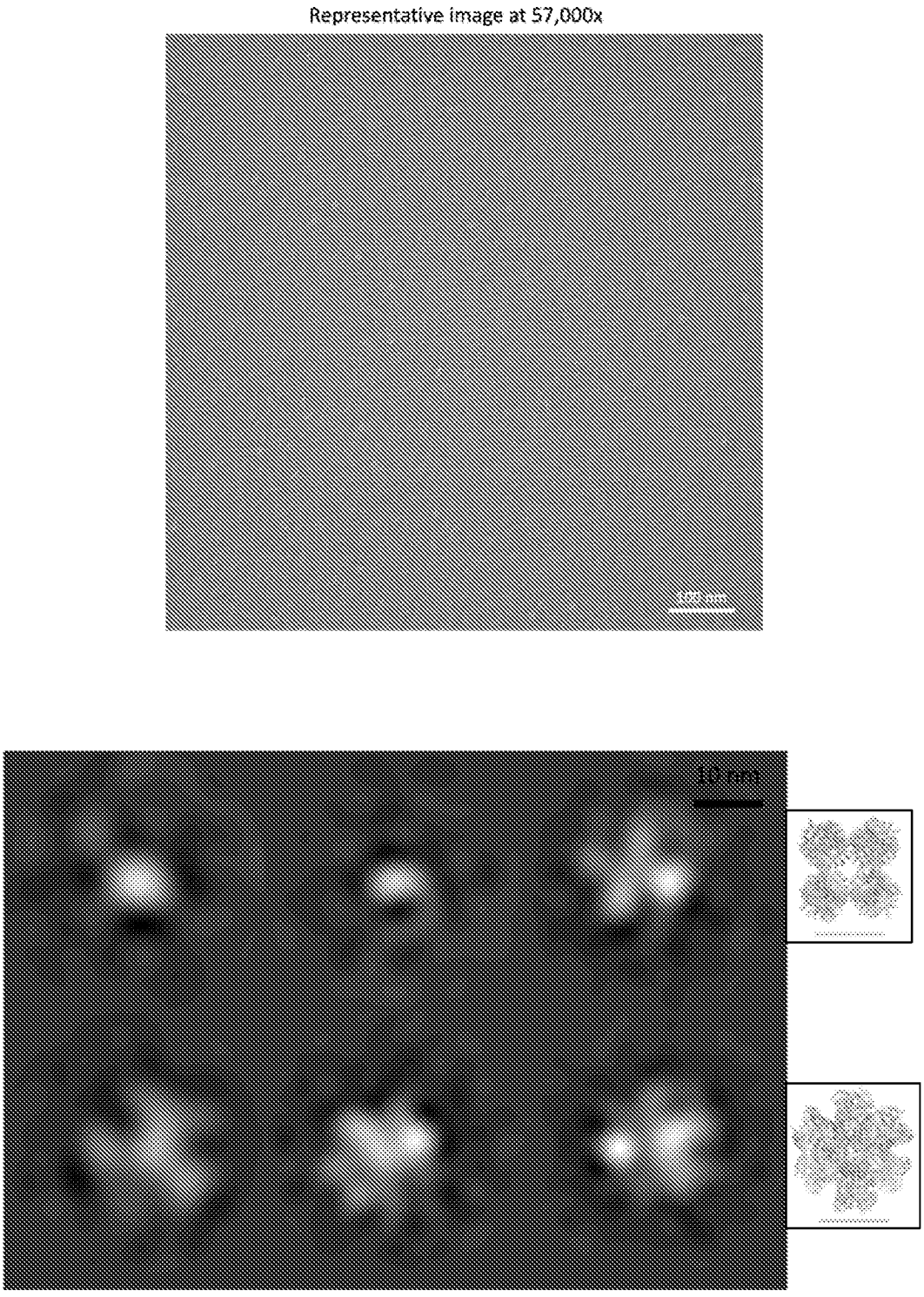
FIG. 4

IgG hinge as linker used for the nanoparticle carrier



LS-hinge2-rTT

FIG. 5



2D classes

HiD-6CCQ-rTT

FIG. 6A

Nanoparticle linked to rTT Carrier by Isopeptide Bond

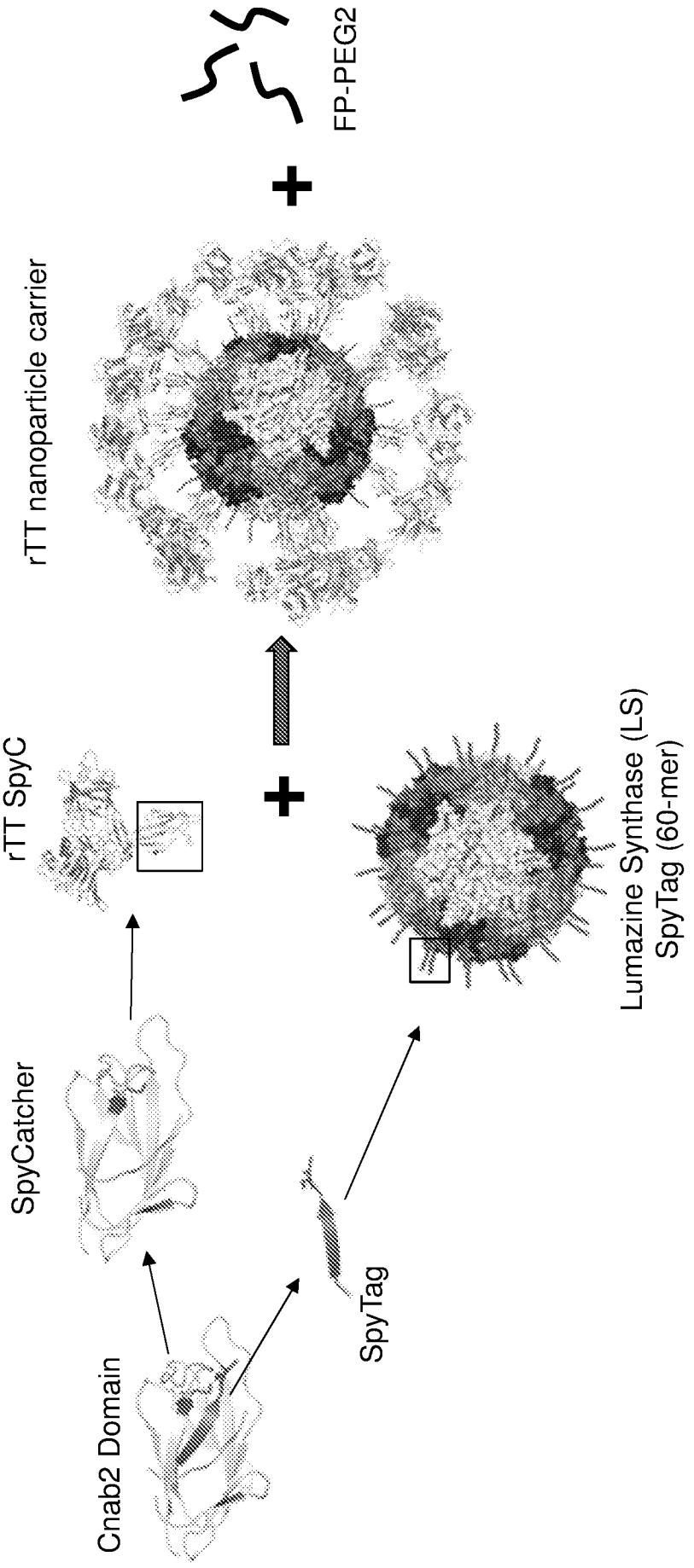


FIG. 6B

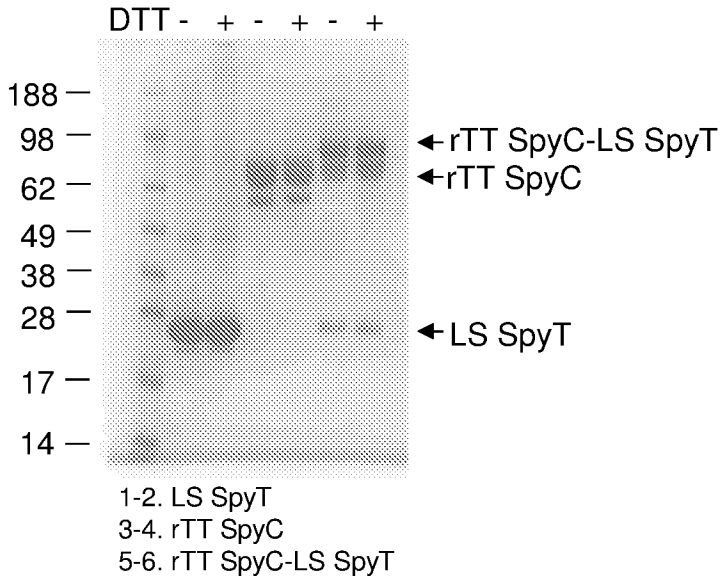


FIG. 6C

LS-SpyT/rTT-SpyC nanoparticle carrier

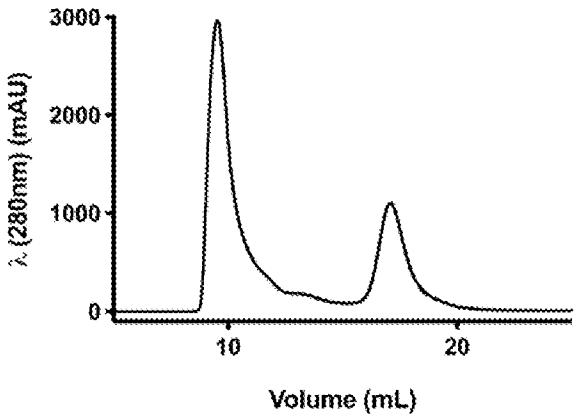


FIG. 7

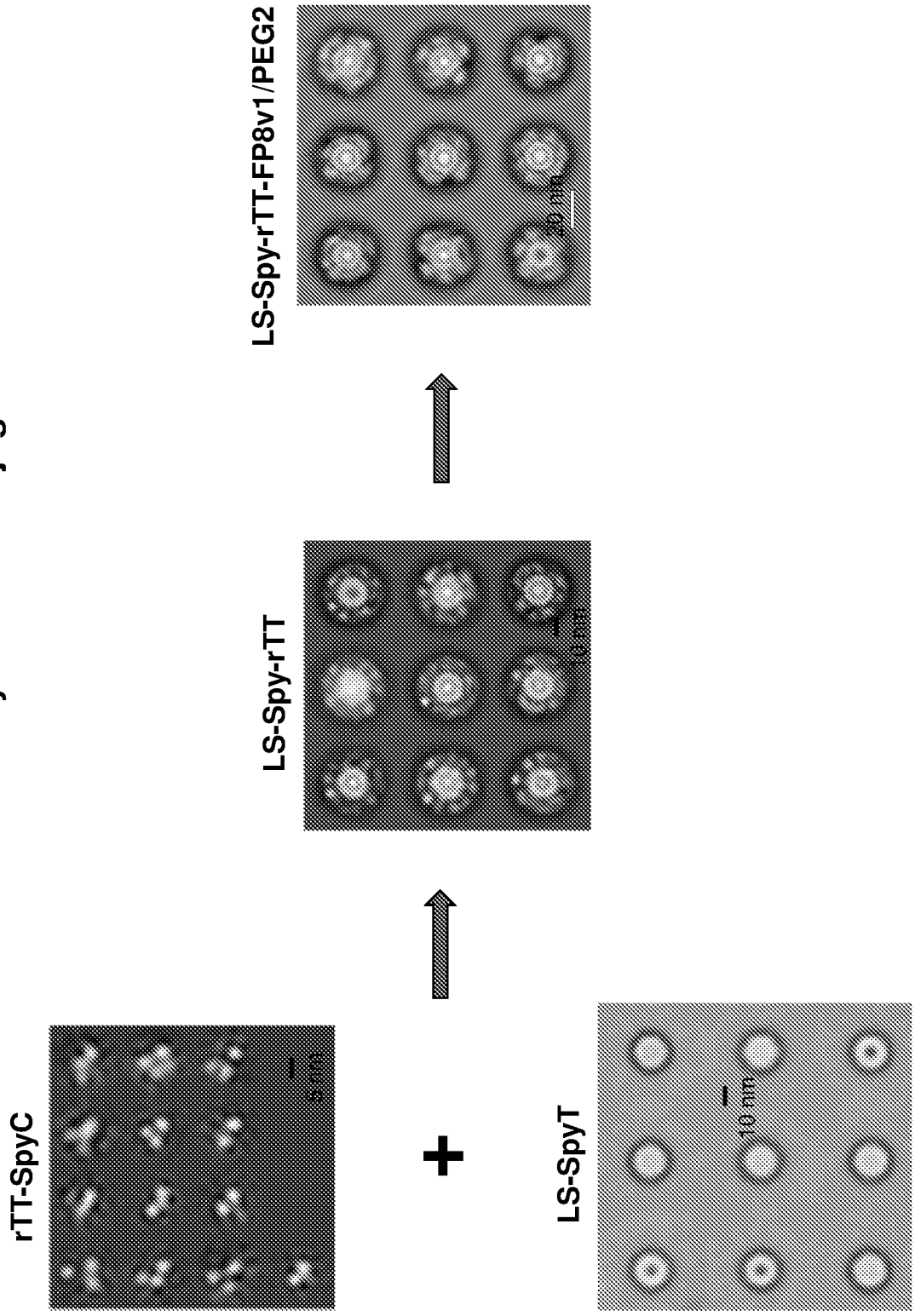
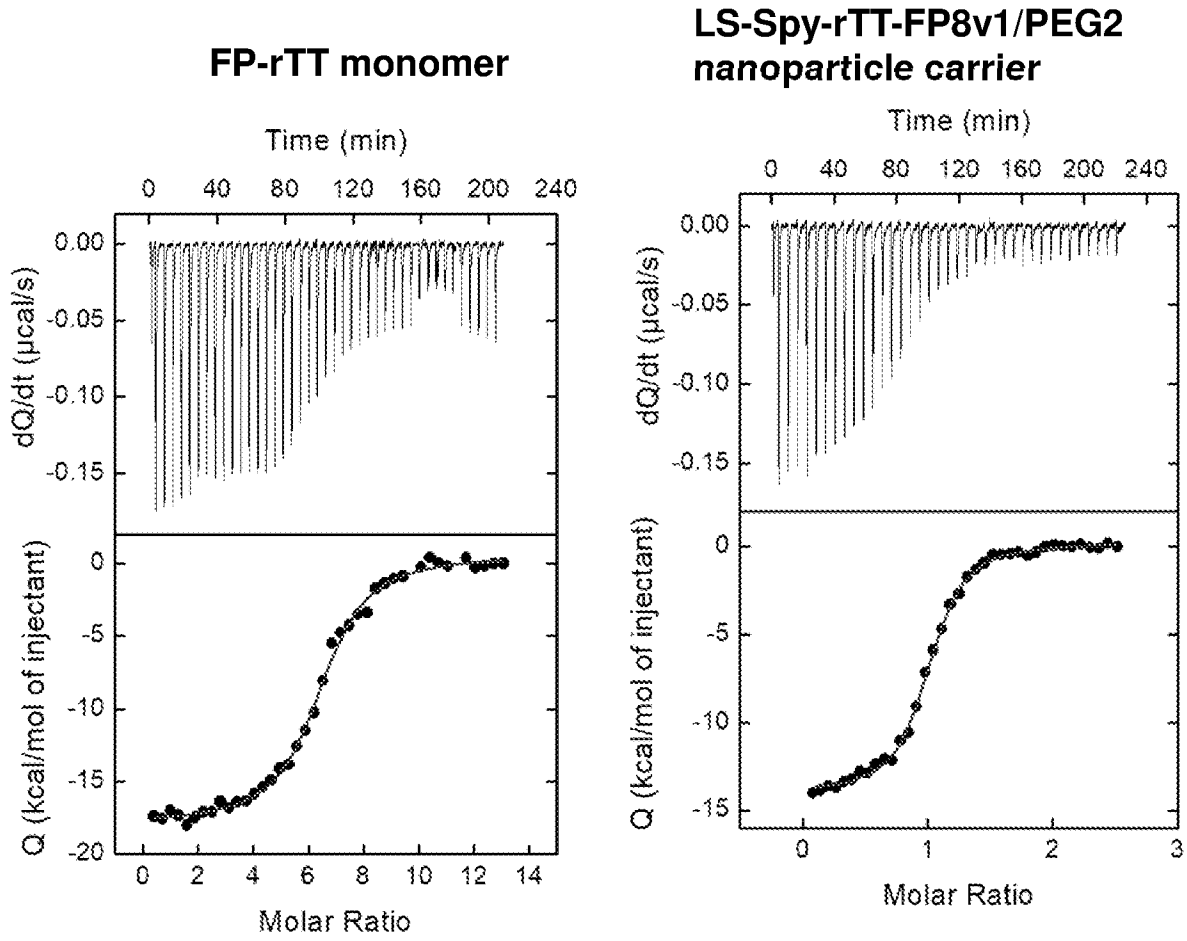

Particle Assembly and FP Conjugation with PEG2 linker
rTT-SpyC
LS-SpyT
LS-Spy-rTT
LS-Spy-rTT-FP8v1/PEG2
50 nm
100 nm
100 nm
100 nm

FIG. 8

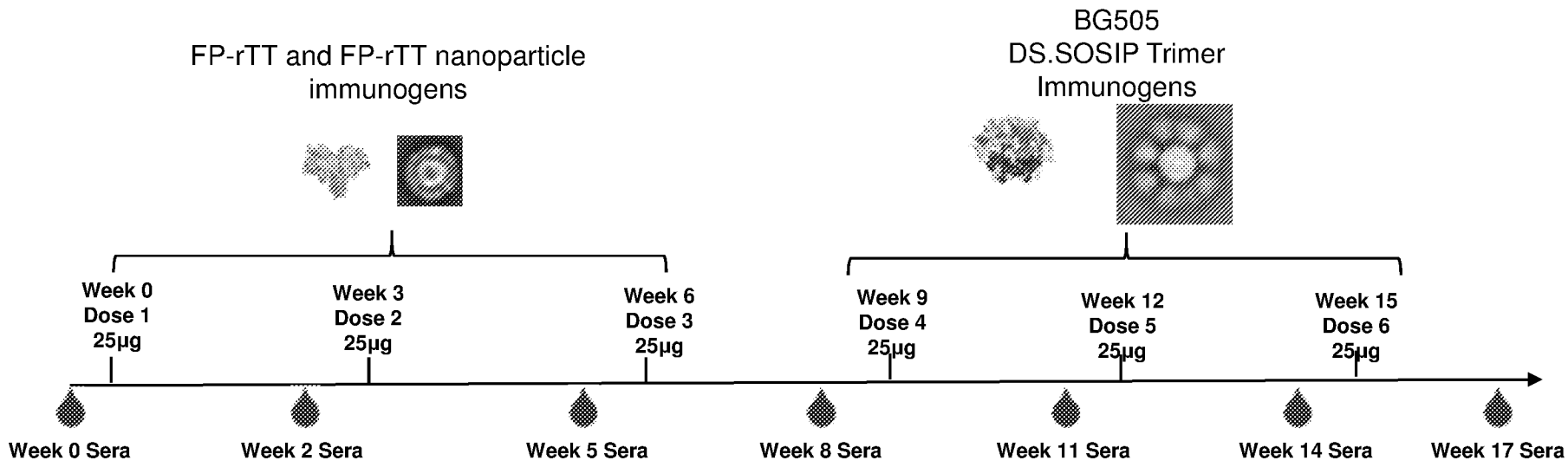
Number of Conjugated FP on Nanoparticle Surface: ITC Analysis



- Each FP-rTT monomer entity has 6 competent VRC34 Fab binding sites
- Each LS-Spy-rTT-FP8v1/PEG2 nanoparticle carrier has 152 – 402 competent VRC34.01 Fab binding sites

FIG. 9

Immunization protocol



Group	# mice (C57 Black/6)	Week 0, 3, and 6 Immunogens	Week 9, 12, and 15 Immunogens
1	10	FP8-rTTHC monomer	BG505 DS-SOSIP trimer
2	10	FP8-rTTHC monomer	BG505 DS-SOSIP trimer on LS nanoparticle
3	10	LS-rTT-FP8 (60mer Nanoparticle)	BG505 DS-SOSIP trimer
4	10	LS-rTT-FP8 (60mer Nanoparticle)	BG505 DS-SOSIP trimer on LS nanoparticle

FIG. 10A

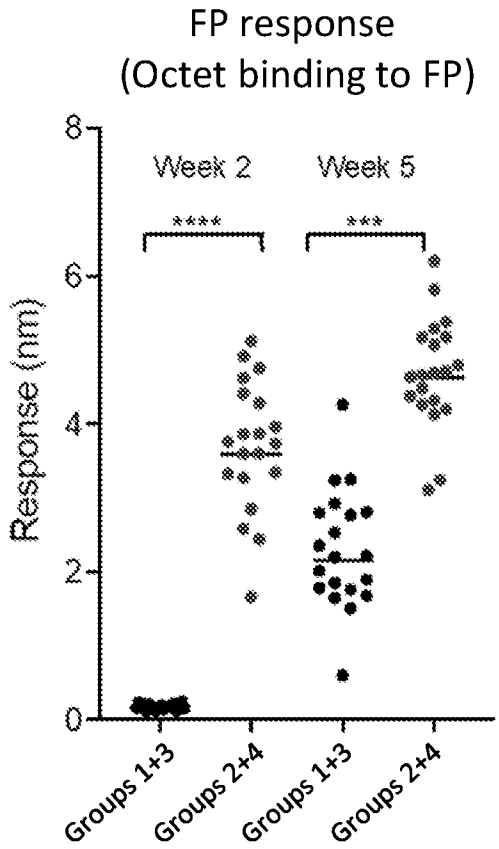


FIG. 10B

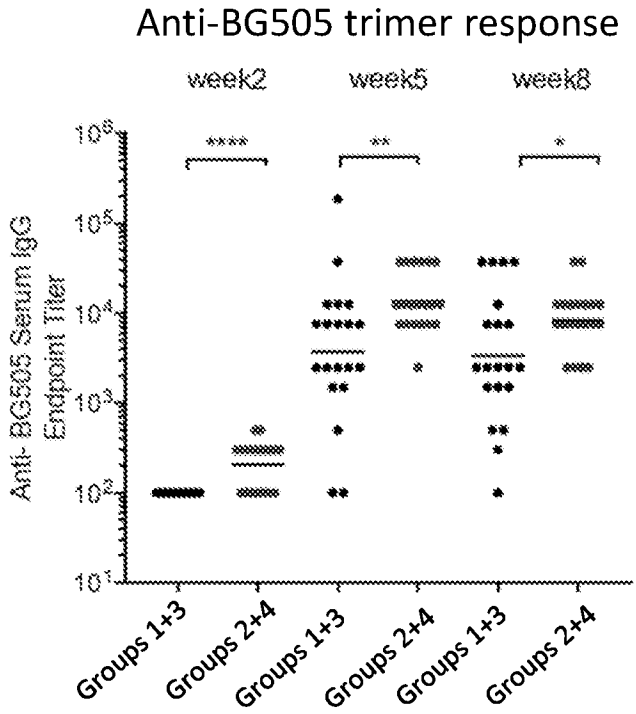


FIG. 10C

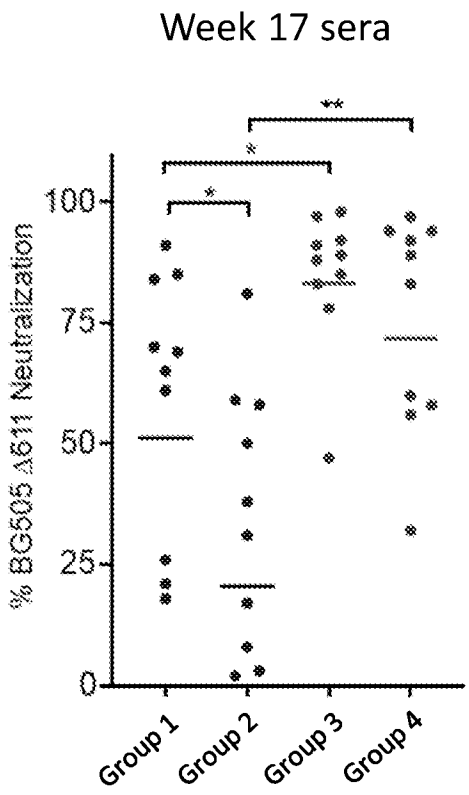

Week 17 sera
% BG505 Δ611 Neutralization
100
75
50
25
0
*
*
**
Group 1
Group 2
Group 3
Group 4

FIG. 11

Disulfide bond stabilized nanoparticle carrier

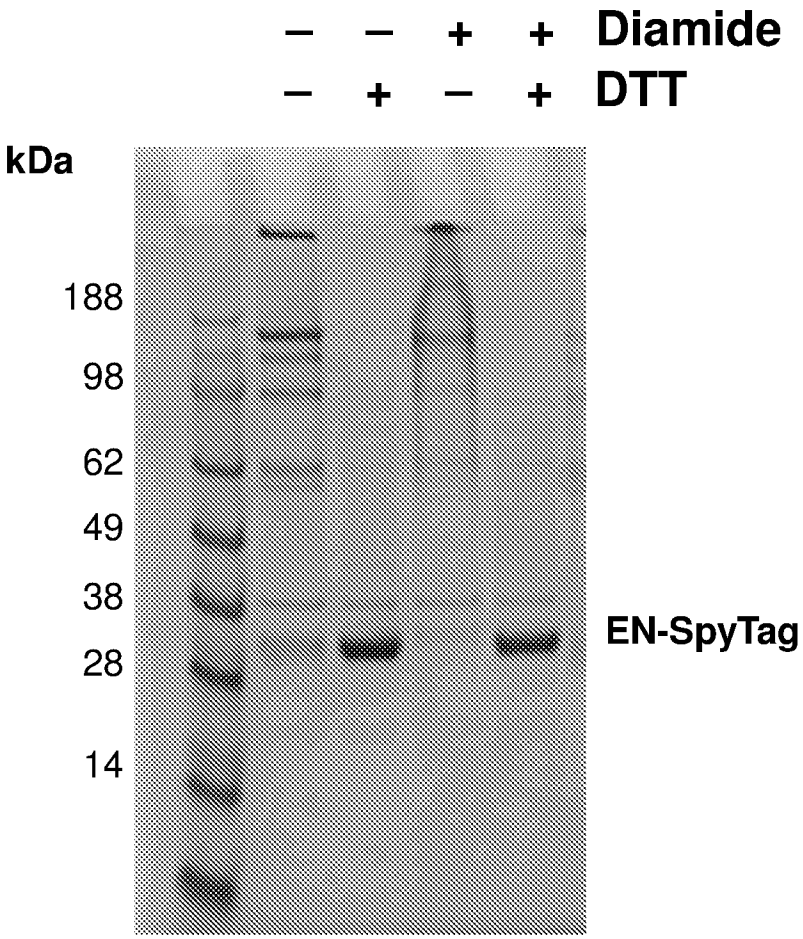


FIG. 12

DS Stabilized Encapsulin Nanoparticle Carrier Is Stable During FP8-Conjugation

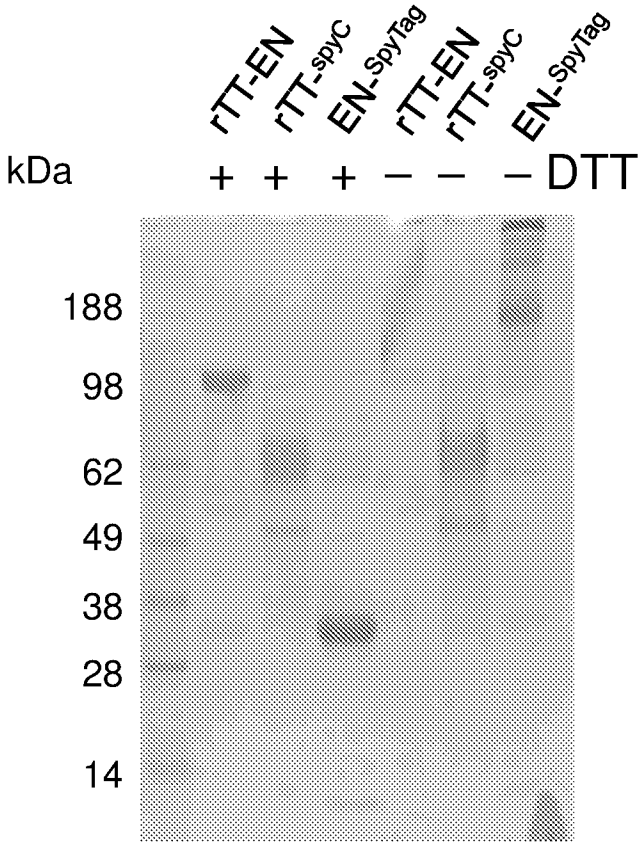
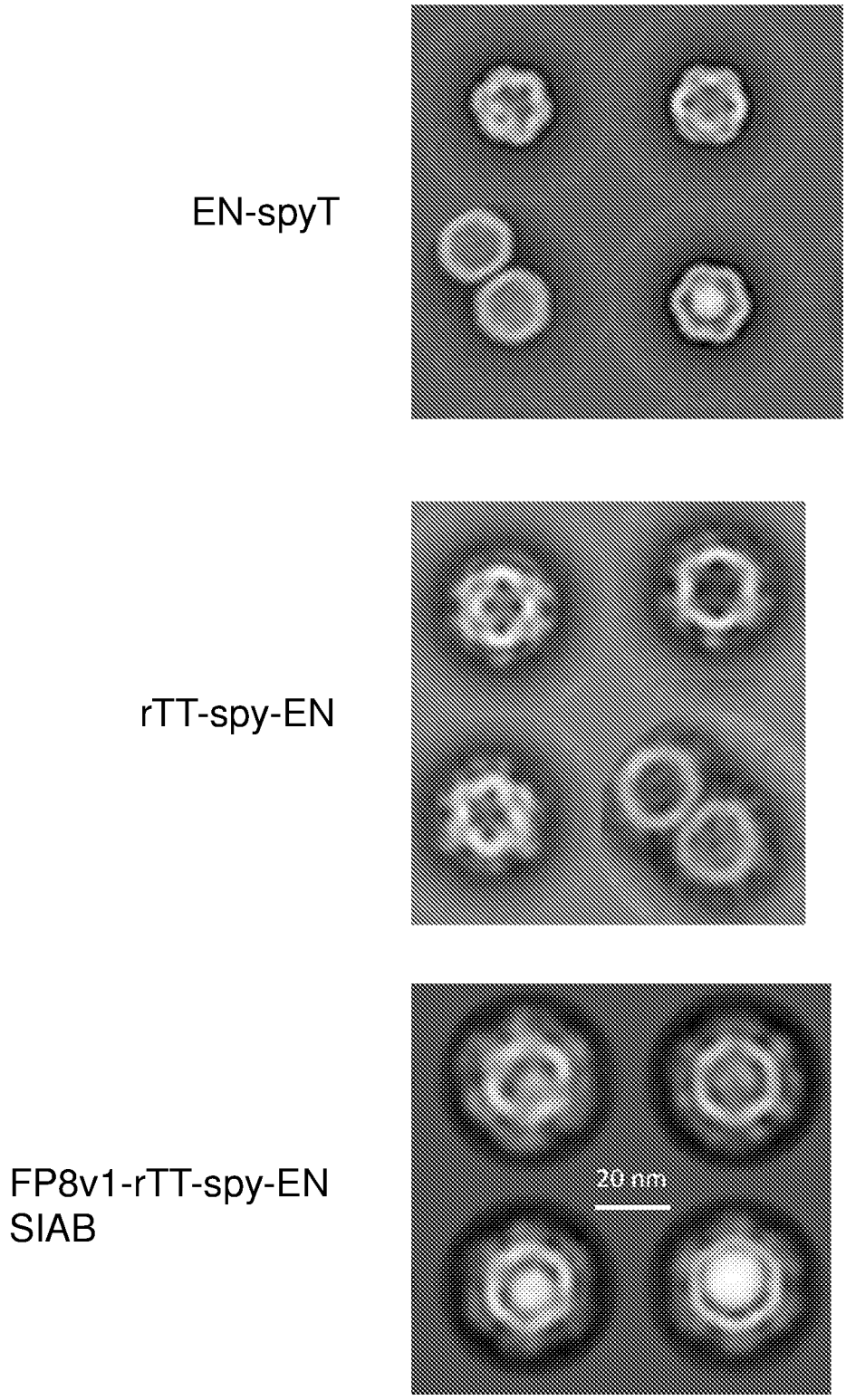


FIG. 13



INTERNATIONAL SEARCH REPORT

International application No
PCT/US2019/052419

A. CLASSIFICATION OF SUBJECT MATTER
INV. A61K39/21 A61K39/385 B82Y5/00 C07K14/08 C07K14/195
C07K14/47 C12N9/10 C07K14/02
ADD.
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
A61K B82Y C07K C12N

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

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, BIOSIS, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 2018/067582 A2 (US HEALTH [US]) 12 April 2018 (2018-04-12) claims 1-56; figures 18-20	1-31
Y	WO 2016/205704 A2 (INT AIDS VACCINE INITIATIVE [US]; SCRIPPS RESEARCH INST [US]) 22 December 2016 (2016-12-22) paragraphs [0091], [0289]; figure 49C	1-31
X,P	WO 2019/089817 A1 (SCRIPPS RESEARCH INST [US]) 9 May 2019 (2019-05-09) claims 1-28; figures 1a,4a	1,5-8, 11, 16-18, 25-31
A	WO 2017/182760 A1 (BIOASTER [FR]) 26 October 2017 (2017-10-26) figures 4-9	1

-/--

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
13 December 2019	24/02/2020

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Lonnoy, Olivier

2

Form PCT/ISA/210 (second sheet) (April 2005)

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JORDAN M FLETCHER ET AL: "Self-assembling cages from coiled-coil peptide modules", SCIENCE, AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE , vol. 340, no. 6132 3 May 2013 (2013-05-03), pages 595-599, XP002726516, ISSN: 0036-8075, DOI: 10.1126/SCIENCE.1233936 Retrieved from the Internet: URL:http://www.sciencemag.org/content/340/6132/595 [retrieved on 2013-04-11] figure 3 -----	1

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2019/052419

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

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

   a. [X] forming part of the international application as filed:

      [X] in the form of an Annex C/ST.25 text file.

      [ ] on paper or in the form of an image file.

   b. [ ] furnished together with the international application under PCT Rule 13*ter*.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.

   c. [ ] furnished subsequent to the international filing date for the purposes of international search only:

      [ ] in the form of an Annex C/ST.25 text file (Rule 13*ter*.1(a)).

      [ ] on paper or in the form of an image file (Rule 13*ter*.1(b) and Administrative Instructions, Section 713).

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

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2019/052419

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

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

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

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

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

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

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

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

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

1-31

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 1-31

An immunogenic conjugate, comprising:a self-assembling protein-nanoparticle carrier comprising a multimer of fusion proteins, wherein each fusion protein comprises a self-assembling protein nanoparticle subunit fused to a heterologous carrier protein, and wherein the fusion proteins self-assemble to form the self-assembling protein-nanoparticle carrier, and HIV-1 Env fusion peptides conjugated to the self-assembling protein-nanoparticle carrier, wherein the HIV-1 Env fusion peptides comprise, from the N-terminus, the amino acid sequence of residue 512 to one of residues 514-521 of a human immunodeficiency virus type 1 (H1V-1) Envelope (Env) protein according to the HXB2 numbering system, and wherein the immunogen elicits an immune response to HIV-1 Env; Method for producing said immunogenic conjugate; Method for generating an immune response to HIV-1 Env in a subject, using said immunogenic conjugate.

---

2. claims: 32-39(partially)

A recombinant self-assembling nanoparticle subunit, comprising: a lumazine synthase nanoparticle subunit comprising cysteine substitutions to introduce one or more non-native disulfide bonds to increase stability of the nanoparticle, wherein the cysteine substitutions comprise 121C and 131C substitutions, 121CG and 131C substitutions, 121GC and 131C substitutions, 7C and 40C substitutions, 3C and 50C substitutions, 82C and 13 ICG substitutions, 5C and 52C substitutions, or 95C and A101C substitutions, or a combination thereof, wherein residue numbering corresponds to a reference lumazine synthase subunit set forth as SEQ ID NO: 25; or a lumazine synthase nanoparticle subunit comprising or consisting of the amino acid sequence set forth as any one of SEQ ID NOs: 306-312; Nucleic acid molecule encoding the recombinant self-assembling nanoparticle subunit; Nanoparticle comprising the recombinant self-assembling nanoparticle subunit.

---

3. claims: 32-39(partially)

A recombinant self-assembling nanoparticle subunit, comprising: an encapsulin nanoparticle subunit comprising cysteine substitutions to introduce one or more non-native disulfide bonds to increase stability of the nanoparticle, wherein the cysteine substitutions comprise 53C and 94C substitutions, 53C and 96C substitutions, or 146C and 185C substitutions, or a combination thereof, wherein residue numbering corresponds to a reference lumazine synthase

subunit set forth as SEQ ID NO: 43; or an encapsulin nanoparticle subunit comprising or consisting of the amino acid sequence set forth as any one of SEQ ID NOs: 313-315; Nucleic acid molecule encoding the recombinant self-assembling nanoparticle subunit; Nanoparticle comprising the recombinant self-assembling nanoparticle subunit.

---

4. claims: 32-39(partially)

A recombinant self-assembling nanoparticle subunit, comprising: an acinetobacter phage AP205 nanoparticle subunit comprising cysteine substitutions to introduce one or more non-native disulfide bonds to increase stability of the nanoparticle, wherein the cysteine substitutions comprise a T81C substitution, 53C and 100C substitution, or 82C and 80C substitutions, or a combination thereof, wherein residue numbering corresponds to a reference lumazine synthase subunit set forth as SEQ ID NO: 316; or a Acinetobacter phage AP205 protein subunit comprising or consisting of the amino acid sequence set forth as any one of SEQ ID NOs: 317-320, Nucleic acid molecule encoding the recombinant self-assembling nanoparticle subunit; Nanoparticle comprising the recombinant self-assembling nanoparticle subunit.

---

5. claims: 32-39(partially)

A recombinant self-assembling nanoparticle subunit, comprising: a Hepatitis B capsid protein nanoparticle subunit comprising cysteine substitutions to introduce one or more non-native disulfide bonds to increase stability of the nanoparticle, wherein the cysteine substitutions comprise 25C and 127C substitutions, 14C and 36C substations, 29C and 127C substitutions, 18C and 36C substitutions, or 29C and 127C substitutions, or a combination thereof, wherein residue numbering corresponds to a reference lumazine synthase subunit set forth as SEQ ID NO: 321; or a Hepatitis B capsid protein subunit comprising or consisting of the amino acid sequence set forth as any one of SEQ ID NOs: 322-326; Nucleic acid molecule encoding the recombinant self-assembling nanoparticle subunit; Nanoparticle comprising the recombinant self-assembling nanoparticle subunit.

---

6. claims: 33-39(partially)

A recombinant self-assembling nanoparticle subunit, comprising: a ferritin nanoparticle subunit comprising or consisting of the amino acid sequence set forth as any one of SEQ ID NOs: 258-305; Nucleic acid molecule encoding the recombinant self-assembling nanoparticle subunit;

Nanoparticle comprising the recombinant self-assembling nanoparticle subunit.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2019/052419

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2018067582 A2	12-04-2018	EP 3518968 A2	07-08-2019
		WO 2018067582 A2	12-04-2018
WO 2016205704 A2	22-12-2016	EP 3334446 A2	20-06-2018
		US 2018194809 A1	12-07-2018
		WO 2016205704 A2	22-12-2016
WO 2019089817 A1	09-05-2019	US 2019125890 A1	02-05-2019
		WO 2019089817 A1	09-05-2019
WO 2017182760 A1	26-10-2017	FR 3050448 A1	27-10-2017
		WO 2017182760 A1	26-10-2017