

CD8-SPECIFIC ANTIBODY CONSTRUCTS AND COMPOSITIONS THEREOF

Related Applications

This application claims priority under 35 U.S.C. § 119(e) to U.S. Provisional Application No. 63/299,254, filed January 13, 2022, and U.S. Provisional Application No. 63/172,518, filed April 8, 2021. The contents of these applications are each incorporated herein by reference in their entirety.

Sequence Listing

The instant application contains a Sequence Listing which has been submitted electronically in ASCII format and is hereby incorporated by reference in its entirety. Said ASCII copy, created on April 6, 2022, is named 15147_0005-00304_SL.txt and is 760 bytes in size.

Field

The present disclosure relates to antibodies or antigen binding fragments thereof that specifically bind human CD8. Also disclosed are fusion proteins comprising a Henipavirus glycoprotein G and a CD8 antibody, or an antigen binding fragment thereof, for targeting and transducing cells expressing CD8. Viral vectors and other compositions containing the fusion proteins, antibodies, or antigen binding fragments thereof, as well as methods of using the fusion proteins, antibodies, or antigen binding fragments thereof are also disclosed.

Background

CD8 (cluster of differentiation 8) is a transmembrane glycoprotein that serves as a co-receptor for the T cell receptor (TCR). CD8 serves multiple functions in immune responses against both both external and internal challenges. In T cells, the CD8 co-receptor functions primarily to bind to a major histocompatibility complex (MHC) molecule to facilitate T cell signaling and aid with cytotoxic T cell antigen interactions. While it is predominantly expressed on the surface of cytotoxic T cells, the CD8 co-receptor can also be found on natural killer cells, cortical thymocytes, and dendritic cells. The CD8 molecule is also used as a marker for cytotoxic T cell population.

There are two isoforms of the CD8 glycoprotein, alpha and beta, and each is encoded by a different gene. To function, CD8 forms a dimer, consisting of a pair of CD8 chains. The most common form of CD8 is composed of a CD8- α and CD8- β chain. Homodimers of the CD8- α chain are also expressed on some cells. The molecular weight of each CD8 chain is about 34 kDa.

T lymphocytes are among the prime targets in gene therapy, even more so since chimeric antigen receptor (CAR) T cells have reached the clinic. Current approaches for T cell engineering mainly rely on *ex vivo* gene transfer methods. Following their isolation from either healthy donors or patients, lymphocytes are activated and subsequently transduced by lentiviral vectors. The modified lymphocytes are then expanded and either used in functional *in vivo* assays or used for *in vivo* applications. *Ex vivo* modification of T lymphocytes, however, has its disadvantages. The complexity of the overall procedure, cost of the manufacturing process, and prolonged *ex vivo* culture negatively impact the quality of the final product. Methods that improve T lymphocyte engineering that use *in vivo* delivery platforms are needed.

In vivo delivery platforms using fusogenic glycoproteins of viral vectors have been shown to be beneficial for targeting, binding, and transducing cells of interest. Certain fusogenic glycoproteins, however, may not be sufficiently stable or expressed on the surface of the viral vector. Thus, improved fusogenic glycoproteins and viral vectors containing those glycoproteins are needed. The provided disclosure addresses this need.

Brief Summary

The present disclosure provides an antibody or antigen binding fragment thereof that specifically binds CD8 α or CD8 β , comprising certain heavy chain complementarity determining regions (HCDR1, HCDR2, and HCDR3) and/or light chain complementarity determining regions (LCDR1, LCDR2, and LCDR3). Another embodiment is an antibody or antigen binding fragment thereof specifically binding CD8 α or CD8 β , comprising certain heavy (VH) and/or light (VL) chain variable regions. The disclosure likewise provides for isolated polynucleotides, vectors, and host cells comprising the anti-CD8 α or CD8 β antibody or antigen binding fragment thereof.

The present disclosure also provides a fusion protein comprising a Henipavirus glycoprotein G (G protein) or a biologically active portion thereof and at least one disclosed CD8 antibody or antigen binding fragment, wherein the antibody or antigen binding fragment is fused to the C-terminus of the G protein or the biologically active portion thereof.

The present disclosure also provides a viral vector comprising a henipavirus F protein molecule or biologically active portion thereof, a henipavirus envelope glycoprotein G (G protein) or a biologically active portion thereof, and at least one disclosed CD8 antibody or antigen binding fragment thereof, wherein the antibody or antigen binding fragment thereof is attached to the C-terminus of the G protein or the biologically active portion thereof.

The present disclosure likewise relates to methods of selectively modulating and transducing CD8⁺ T cells using the disclosed viral vectors. Also disclosed are methods of delivering an exogenous agent to a subject, comprising administering to the subject the disclosed viral vectors, in which the viral vector further comprises an exogenous agent. The present disclosure also relates to methods of treating cancer in a subject, comprising administering to the subject the disclosed viral vectors.

Brief Description of Drawings

Fig. 1A–1B shows the transduction efficiencies of the disclosed anti-CD8 antibodies in several human cell lines and a cell line expressing M. nemestrina CD8 α and CD8 β .

Fig. 2 shows the titers of the disclosed anti-CD8 antibodies on SupT1 cells.

Fig. 3 shows the titers of the disclosed anti-CD8 antibodies on cells expressing M. nemestrina CD8 α and CD8 β .

Fig. 4 shows the titers of select CD8 binders on CD8 α only or CD8 $\alpha\beta$ overexpressing cells.

Fig. 5 shows the transduction efficiencies of a CD8 scFv (CD8 binder 52) of the disclosure on human or M. nemestrina (NHP) peripheral blood mononuclear cells (PBMCs).

Fig. 6 shows the transduction efficiencies of a CD8 scFv (CD8 binder 46) of the disclosure on human or M. nemestrina (NHP) PBMCs.

Figs. 7A–7D shows the transduction effects of a Nipah G pseudotyped viral vector attached to CD8 binder 1 (**Fig. 7A**), CD8 binder 5 (**Fig. 7B**), CD8 binder 46 (**Fig. 7C**), and CD8 binder 52 (**Fig. 7D**) of the disclosure on human and nemestrina PBMCs (human donor 905C and nemestrina donor 24).

Figs. 8A–8B show the average titers (**Fig. 8A**) and titers by donor (905C, 703C, or 2101C) (**Fig. 8B**) of several CD8 scFvs of the disclosure on human PBMCs.

Figs. 9A–9B depict the transduction efficiencies of the CD8 binders of the disclosure on human PBMC in transgenic mice *in vivo* using CD8+ (**Fig. 9A**) and CD8- (**Fig. 9B**) cells.

Figs. 10A–10B depict tumor killing *in vitro* using CD8 fusogen-generated CD19CAR T cells.

Figs. 11A–11C depict tumor growth over time following injection of human peripheral blood mononuclear cells (hPBMC) (**Fig. 11A**), with prior T cell activation with CD3/CD28 complexes, and then injection a day after of CD8-CD19CAR LV, and following injection of non-activated hPBMC (without prior T cell activation), and then injection a day after of CD8-CD19CAR LV (**Fig. 11B**). **FIG. 11C** shows the percent of CD8⁺CD19CAR⁺ cells in total recovered live lymphocytes from spleen, bone marrow or peripheral blood following injection of CD8-CD19CAR LV, as indicated in top right quadrant of the FACs plots in both PBMC control (top plots) and CD8 fusosome-treated animals (bottom plots).

Fig. 12 depicts B cell levels in a non-human primate (NHP) model following administration of a lentiviral vector pseudotyped with an anti-CD8 binding protein targeting CD8⁺ T cells to deliver a CD20 CAR transgene.

Fig. 13 depicts normalized B cell counts in non-human primates following administration of vehicle (Control 1-2) or a lentiviral vector pseudotyped with an anti-CD8 binding protein targeting CD8⁺ T cells to deliver a CD20 CAR transgene (Treated 1-6).

Fig. 14 depicts an exemplary system for administration of a lentiviral vector comprising a CD8 binding agent to a subject.

Detailed Description

Unless defined otherwise, all terms of art, notations, and other technical and scientific terms or terminology used herein are intended to have the same meaning

as is commonly understood by one of ordinary skill in the art to which the claimed subject matter pertains. In some cases, terms with commonly understood meanings are defined herein for clarity and/or for ready reference, and the inclusion of such definitions herein should not necessarily be construed to represent a substantial difference over what is generally understood in the art.

Unless defined otherwise, all technical and scientific terms, acronyms, and abbreviations used herein have the same meaning as commonly understood by one of ordinary skill in the art to which the disclosure pertains. Unless indicated otherwise, abbreviations and symbols for chemical and biochemical names is per IUPAC-IUB nomenclature. Unless indicated otherwise, all numerical ranges are inclusive of the values defining the range as well as all integer values in-between. As used herein, the articles “a” and “an” refer to one or to more than one (i.e. to at least one) of the grammatical object of the article. By way of example, “an element” means one element or more than one element.

As used herein, the term “about” will be understood by persons of ordinary skill in the art and will vary to some extent on the context in which it is used. In some embodiments, the term “about” when referring to a measurable value such as an amount, a temporal duration, and the like, is meant to encompass art-accepted variations based on standard errors in making such measurements. In some embodiments, the term “about” when referring to such values, is meant to encompass variations of $\pm 20\%$ or $\pm 10\%$, more preferably $\pm 5\%$, even more preferably $\pm 1\%$, and still more preferably $\pm 0.1\%$ from the specified value, as such variations are appropriate to perform the disclosed methods.

As used herein, “CD8” or “cluster of differentiation 8” refers to a transmembrane glycoprotein which is a specific marker for a subclass of T cells (which includes cytotoxic T cells). CD8 assembles as either a heterodimer of the CD8 alpha (“CD8 α ” or “CD8A”) and CD8 beta (“CD8 β ” or “CD8B”) subunits (“CD8 $\alpha\beta$ ” or “CD8AB”), or a CD8 alpha homodimer (“CD8 $\alpha\alpha$ ” or “CD8AA”). The assembled dimeric CD8 complex acts as a co-receptor together with the T cell receptor (TCR) to recognize antigen presentation by MHC class I cells. CD8 plays a role in the development of T cells and activation of mature T cells.

As used herein, “affinity” refers to the strength of the sum total of noncovalent interactions between a single binding site of a molecule (e.g., an antibody) and its

binding partner (e.g., an antigen). The affinity of a molecule for its partner can generally be represented by the equilibrium dissociation constant (K_D) (or its inverse equilibrium association constant, K_A). Affinity can be measured by common methods known in the art, including those described herein. See, for example, Pope M.E., Soste M.V., Eyford B.A., Anderson N.L., Pearson T.W., (2009) *J. Immunol. Methods*. 341(1-2):86-96 and methods described therein.

As used herein, "antibody" is meant in a broad sense and includes immunoglobulin molecules including monoclonal antibodies including murine, human, humanized and chimeric antibodies, antibody fragments, bispecific or multispecific antibodies formed from at least two intact antibodies or antibody fragments, dimeric, tetrameric or multimeric antibodies, single chain antibodies, and any other modified configuration of the immunoglobulin molecule that comprises an antigen recognition site of the required specificity.

Immunoglobulins can be assigned to five major classes, namely IgA, IgD, IgE, IgG, and IgM, depending on the heavy chain constant domain amino acid sequence. IgA and IgG are further sub-classified to IgA1, IgA2, IgG1, IgG2, IgG3, and IgG4.

Antibody light chains of any vertebrate species can be assigned to one of two clearly distinct types, namely kappa (K) and lambda (A), based on the amino acid sequences of their constant domains.

As used herein, "antigen binding fragment" or "antibody fragment" refers to a portion of an immunoglobulin molecule that retains the heavy chain and/or the light chain antigen binding site, such as a heavy chain complementarity determining regions (HCDR) 1 (HCDR1), 2 (HCDR2), and 3 (HCDR3), a light chain complementarity determining regions (LCDR) 1 (LCDR1), 2 (LCDR2), and 3 (LCDR3), a heavy chain variable region (VH), or a light chain variable region (VL). Antibody fragments include a Fab fragment (a monovalent fragment consisting of the VL or the VH); a F(ab) 2 fragment (a bivalent fragment comprising two Fab fragments linked by a disulfide bridge at the hinge region); a Fd fragment consisting of the VH and CH1 domains; a Fv fragment consisting of the VL and VH domains of a single arm of an antibody; a dAb fragment, which consists of a VH domain; and a variable domain (VHH) from, e.g., human or camelid origin. VH and VL domains can be engineered and linked together via a synthetic linker to form various types of single chain antibody designs in which the VH/VL domains pair intramolecularly, or intermolecularly in those cases

in which the VH and VL domains are expressed by separate single chain antibody constructs, to form a monovalent antigen binding site, such as a single-chain Fv (scFv) or diabody. These antibody fragments are obtained using well known techniques and the fragments are characterized in the same manner as are intact antibodies.

An antibody variable region consists of a “framework” region interrupted by three “antigen binding sites.” The antigen binding sites are defined using various terms, including, for example (i) “Complementarity Determining Regions” (CDRs), three in the VH (HCDR1, HCDR2, HCDR3) and three in the VL (LCDR1, LCDR2, LCDR3) (Wu and Kabat, *J Exp Med* 132:211-50, 1970; Kabat et al., *Sequences of Proteins of Immunological Interest*, 5th Ed. Public Health Service, National Institutes of Health, Bethesda, Md., 1991), and (ii) “Hypervariable regions,” “HVR,” or “HV,” three in the VH (H1, H2, H3) and three in the VL (L1, L2, L3) (Chothia and Lesk *Mol Biol* 196:901-17, 1987). Other terms include “IMGT-CDRs” (Lefranc et al., *Dev Comparat Immunol* 27:55-77, 2003) and “Specificity Determining Residue Usage” (SDRU) (Almagro *Mol Recognit*, 17:132-43, 2004). The International ImMunoGeneTics (IMGT) database ([http://www_imgt.org](http://www.imgt.org)) provides a standardized numbering and definition of antigen-binding sites. The correspondence between CDRs, HVs, and IMGT delineations is described in Lefranc et al., *Dev Comparat Immunol* 27:55-77, 2003.

The term “framework,” or “FR” or “framework sequence” refers to the remaining sequences of a variable region other than those sequences defined to be antigen binding sites. Because the antigen binding site can be defined by various terms as described above, the exact amino acid sequence of a framework depends on how the antigen-binding site was defined.

The term “CDR” denotes a complementarity determining region as defined by at least one manner of identification to one of skill in the art. The precise amino acid sequence boundaries of a given CDR or FR can be readily determined using any of a number of well-known schemes, including those described by Kabat et al. (1991), “Sequences of Proteins of Immunological Interest,” 5th Ed. Public Health Service, National Institutes of Health, Bethesda, MD (“Kabat” numbering scheme); Al-Lazikani et al., (1997) *JMB* 273,927-948 (“Chothia” numbering scheme); MacCallum et al., *J. Mol. Biol.* 262:732-745 (1996), “Antibody-antigen interactions: Contact analysis and

binding site topography," J. Mol. Biol. 262, 732-745." ("Contact" numbering scheme); Lefranc MP *et al.*, "IMGT unique numbering for immunoglobulin and T cell receptor variable domains and Ig superfamily V-like domains," Dev Comp Immunol, 2003 Jan;27(1):55-77 ("IMGT" numbering scheme); Honegger A and Plückthun A, "Yet another numbering scheme for immunoglobulin variable domains: an automatic modeling and analysis tool," J Mol Biol, 2001 Jun 8;309(3):657-70, ("Aho" numbering scheme); and Martin *et al.*, "Modeling antibody hypervariable loops: a combined algorithm," PNAS, 1989, 86(23):9268-9272, ("AbM" numbering scheme).

The boundaries of a given CDR or FR may vary depending on the scheme used for identification. For example, the Kabat scheme is based on structural alignments, while the Chothia scheme is based on structural information. Numbering for both the Kabat and Chothia schemes is based upon the most common antibody region sequence lengths, with insertions accommodated by insertion letters, for example, "30a," and deletions appearing in some antibodies. The two schemes place certain insertions and deletions ("indels") at different positions, resulting in differential numbering. The Contact scheme is based on analysis of complex crystal structures and is similar in many respects to the Chothia numbering scheme. The AbM scheme is a compromise between Kabat and Chothia definitions based on that used by Oxford Molecular's AbM antibody modeling software.

In some embodiments, CDRs can be defined in accordance with any of the Chothia numbering schemes, the Kabat numbering scheme, the IMGT numbering scheme, a combination of Kabat, IMGT, and Chothia, the AbM definition, and/or the contact definition. A sdAb variable domain comprises three CDRs, designated CDR1, CDR2, and CDR3. **Table 1**, below, lists exemplary position boundaries of CDR-H1, CDR-H2, CDR-H3 as identified by Kabat, Chothia, AbM, and Contact schemes, respectively. For CDR-H1, residue numbering is listed using both the Kabat and Chothia numbering schemes. FRs are located between CDRs, for example, with FR-H1 located before CDR-H1, FR-H2 located between CDR-H1 and CDR-H2, FR-H3 located between CDR-H2 and CDR-H3 and so forth. It is noted that because the shown Kabat numbering scheme places insertions at H35A and H35B, the end of the Chothia CDR-H1 loop when numbered using the shown Kabat numbering convention varies between H32 and H34, depending on the length of the loop.

Table 1. Boundaries of CDRs according to various numbering schemes.				
CDR	Kabat	Chothia	AbM	Contact
CDR-H1 (Kabat Numbering ¹)	H31--H35B	H26--H32..34	H26--H35B	H30--H35B
CDR-H1 (Chothia Numbering ²)	H31--H35	H26--H32	H26--H35	H30--H35
CDR-H2	H50--H65	H52--H56	H50--H58	H47--H58
CDR-H3	H95--H102	H95--H102	H95--H102	H93--H101

¹ - Kabat *et al.* (1991), "Sequences of Proteins of Immunological Interest," 5th Ed. Public Health Service, National Institutes of Health, Bethesda, MD

² - Al-Lazikani *et al.*, (1997) JMB 273,927-948.

Thus, unless otherwise specified, a "CDR" or "complementary determining region," or individual specified CDRs (e.g., CDR-H1, CDR-H2, CDR-H3), of a given antibody or region thereof, such as a variable region thereof, should be understood to encompass a (or the specific) complementary determining region as defined by any of the aforementioned schemes. For example, where it is stated that a particular CDR (e.g., a CDR-H3) contains the amino acid sequence of a corresponding CDR in a given sdAb amino acid sequence, it is understood that such a CDR has a sequence of the corresponding CDR (e.g., CDR-H3) within the sdAb, as defined by any of the aforementioned schemes. It is understood that any antibody, such as a sdAb, includes CDRs and such can be identified according to any of the other aforementioned numbering schemes or other numbering schemes known to a skilled artisan.

As used herein, "Fv" refers to the minimum antibody fragment which contains a complete antigen-recognition and antigen-binding site. This region consists of a dimer of one heavy chain and one light chain variable domain in tight, non-covalent association. It is in this configuration that the three hypervariable regions of each variable domain interact to define an antigen-binding site on the surface of the VH-VL dimer. Collectively, the six hypervariable regions confer antigen-binding specificity to the antibody. However, even a single variable domain (or half of an Fv comprising only three hypervariable regions specific for an antigen) may have the ability to recognize and bind an antigen, although at a lower affinity than the entire binding site.

As used herein, “single-chain Fv” or “scFv” antibody fragments comprise the VH and VL domains of an antibody, wherein these domains are present in a single polypeptide chain. Preferably, the Fv polypeptide further comprises a polypeptide linker between the VH and VL domains which enables the scFv to form the desired structure for antigen binding. For a review of scFv see Pluckthun in *The Pharmacology of Monoclonal Antibodies*, vol. 113, Rosenberg and Moore eds., Springer-Verlag, New York, pp. 269-315 (1994).

As used herein, “VHH” or “VHH antibodies” refer to single domain antibodies that consist of the variable region of a heavy chain of an IgG antibody. For example, the terms “VHH” and “VHH antibody” can refer to the antigen binding domain of a heavy chain IgG (hclgG) molecule produced by a Camelidae family mammal (e.g., llamas, camels, and alpacas).

As used herein, the term “specifically binds” to a target molecule, such as an antigen, means that a binding molecule, such as a single domain antibody, reacts or associates more frequently, more rapidly, with greater duration, and/or with greater affinity with a particular target molecule than it does with alternative molecules. A binding molecule, such as a sdAb or scFv, “specifically binds” to a target molecule if it binds with greater affinity, avidity, more readily, and/or with greater duration than it binds to other molecules. It is understood that a binding molecule, such as a sdAb or scFv, that specifically binds to a first target may or may not specifically bind to a second target. As such, “specific binding” does not necessarily require (although it can include) exclusive binding.

As used herein, “percent (%) amino acid sequence identity” and “homology” with respect to a peptide, polypeptide or antibody sequence are used interchangeably and are defined as the percentage of amino acid residues in a candidate sequence that are identical with the amino acid residues in another peptide or polypeptide sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Alignment for purposes of determining percent amino acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN, or MEGALIGN (DNASTAR) software. Those skilled in the art can determine appropriate parameters for measuring

alignment, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared.

An amino acid substitution may include but is not limited to the replacement of one amino acid in a polypeptide with another amino acid. Exemplary substitutions are shown in **Table 2**. Amino acid substitutions may be introduced into an antibody of interest and the products screened for a desired activity, for example, retained/improved binding.

Table 2

Original Residue	Exemplary Substitutions
Ala (A)	Val; Leu; Ile
Arg (R)	Lys; Gln; Asn
Asn (N)	Gln; His; Asp; Lys; Arg
Asp (D)	Glu; Asn
Cys (C)	Ser; Ala
Gln (Q)	Asn; Glu
Glu (E)	Asp; Gln
Gly (G)	Ala
His (H)	Asn; Gln; Lys; Arg
Ile (I)	Leu; Val; Met; Ala; Phe; Norleucine
Leu (L)	Norleucine; Ile; Val; Met; Ala; Phe
Lys (K)	Arg; Gln; Asn
Met (M)	Leu; Phe; Ile
Phe (F)	Trp; Leu; Val; Ile; Ala; Tyr
Pro (P)	Ala
Ser (S)	Thr
Thr (T)	Val; Ser
Trp (W)	Tyr; Phe
Tyr (Y)	Trp; Phe; Thr; Ser
Val (V)	Ile; Leu; Met; Phe; Ala; Norleucine

Amino acids may be grouped according to common side-chain properties:

(1) hydrophobic: Norleucine, Met, Ala, Val, Leu, Ile;

- (2) neutral hydrophilic: Cys, Ser, Thr, Asn, Gln;
- (3) acidic: Asp, Glu;
- (4) basic: His, Lys, Arg;
- (5) residues that influence chain orientation: Gly, Pro;
- (6) aromatic: Trp, Tyr, Phe.

Non-conservative substitutions will entail exchanging a member of one of these classes for another class. The term, “corresponding to” with reference to nucleotide or amino acid positions of a sequence, such as set forth in the Sequence Listing, refers to nucleotides or amino acid positions identified upon alignment with a target sequence based on structural sequence alignment or using a standard alignment algorithm, such as the GAP algorithm. For example, corresponding residues of a similar sequence (e.g. fragment or species variant) can be determined by alignment to a reference sequence by structural alignment methods. By aligning the sequences, one skilled in the art can identify corresponding residues, for example, using conserved and identical amino acid residues as guides.

The term “isolated” as used herein refers to a molecule that has been separated from at least some of the components with which it is typically found in nature or produced. For example, a polypeptide is referred to as “isolated” when it is separated from at least some of the components of the cell in which it was produced. When a polypeptide is secreted by a cell after expression, physically separating the supernatant containing the polypeptide from the cell that produced it is considered to be “isolating” the polypeptide. Similarly, a polynucleotide is referred to as “isolated” when it is not part of the larger polynucleotide (such as, for example, genomic DNA or mitochondrial DNA, in the case of a DNA polynucleotide) in which it is typically found in nature, or is separated from at least some of the components of the cell in which it was produced. Thus, a DNA polynucleotide that is contained in a vector inside a host cell may be referred to as “isolated.”

As used herein, “lipid particle” refers to any biological or synthetic particle that contains a bilayer of amphipathic lipids enclosing a lumen or cavity. Typically, a lipid particle does not contain a nucleus. Examples of lipid particles include nanoparticles, viral-derived particles, or cell-derived particles. Such lipid particles include, but are not limited to, viral particles (e.g. lentiviral particles), virus-like particles, viral vectors (e.g., lentiviral vectors), exosomes, enucleated cells, vesicles (e.g., microvesicles,

membrane vesicles, extracellular membrane vesicles, plasma membrane vesicles, and giant plasma membrane vesicles), apoptotic bodies, mitoparticles, pyrenocytes, or lysosomes. In some embodiments, a lipid particle can be a fusosome. In some embodiments, the lipid particle is not a platelet.

As used herein a “biologically active portion,” such as with reference to a protein such as a G protein or an F protein, refers to a portion of the protein that exhibits or retains an activity or property of the full-length of the protein. For example, a biologically active portion of an F protein retains fusogenic activity in conjunction with the G protein when each are embedded in a lipid bilayer. A biologically active portion of the G protein retains fusogenic activity in conjunction with an F protein when each is embedded in a lipid bilayer. The retained activity can include 10%-150% or more of the activity of a full-length or wild-type F protein or G protein. Examples of biologically active portions of F and G proteins include truncations of the cytoplasmic domain, e.g. truncations of up to 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 22, 25, 30, 33, 34, 35, or more contiguous amino acids, see e.g. Khetawat and Broder 2010 Virology Journal 7:312; Witting et al. 2013 Gene Therapy 20:997-1005; published international; patent application No. WO/2013/148327.

As used herein, “G protein” refers to a henipavirus envelope attachment glycoprotein G or biologically active portion thereof. “F protein” refers to a henipavirus fusion protein F or biologically active portion thereof. The F and G proteins may be from a Hendra (HeV) or a Nipah (NiV) virus, and may be a wild-type protein or may be a variant thereof that exhibits reduced binding for the native binding partner. The F (fusion) and G (attachment) glycoproteins mediate cellular entry of Nipah virus. The G protein initiates infection by binding to the cellular surface receptor ephrin-B2 (EphB2) or EphB3. The subsequent release of the viral genome into the cytoplasm is mediated by the action of the F protein, which induces the fusion of the viral envelope with cellular membranes. The efficiency of transduction of targeted lipid particles can be improved by engineering hyperfusogenic mutations in one or both of the F protein (such as NiV-F) and G protein (such as NiV-G).

As used herein, “fusosome” refers to a particle containing a bilayer of amphipathic lipids enclosing a lumen or cavity and a fusogen that interacts with the amphipathic lipid bilayer. In some embodiments, the fusosome comprises a nucleic acid. In some embodiments, the fusosome is a membrane enclosed preparation. In some

embodiments, the fusosome is derived from a source cell. As used herein, “fusosome composition” refers to a composition comprising one or more fusosomes. As used herein, “fusogen” refers to an agent or molecule that creates an interaction between two membrane enclosed lumens. In embodiments, the fusogen facilitates fusion of the membranes. In other embodiments, the fusogen creates a connection, e.g., a pore, between two lumens (e.g., a lumen of a retroviral vector and a cytoplasm of a target cell). In some embodiments, the fusogen comprises a complex of two or more proteins, e.g., wherein neither protein has fusogenic activity alone. In some embodiments, the fusogen comprises a targeting domain.

As used herein, a “re-targeted fusogen” refers to a fusogen that comprises a targeting moiety having a sequence that is not part of the naturally-occurring form of the fusogen. In embodiments, the fusogen comprises a different targeting moiety relative to the targeting moiety in the naturally-occurring form of the fusogen. In embodiments, the naturally-occurring form of the fusogen lacks a targeting domain, and the re-targeted fusogen comprises a targeting moiety that is absent from the naturally-occurring form of the fusogen. In embodiments, the fusogen is modified to comprise a targeting moiety. In embodiments, the fusogen comprises one or more sequence alterations outside of the targeting moiety relative to the naturally-occurring form of the fusogen, e.g., in a transmembrane domain, fusogenically active domain, or cytoplasmic domain.

As used herein, a “targeted envelope protein” refers to a polypeptide that contains a henipavirus G protein (G protein) attached to a single domain antibody (sdAb) variable domain, such as a VL or VH sdAb, a scFv, a nanobody, a camelid VHH domain, a shark IgNAR, or fragments thereof, that target a molecule on a desired cell type. In some such embodiments, the attachment may be directly or indirectly via a linker, such as a peptide linker. The “targeted envelope protein” may also be referred to as a “fusion protein” comprising the G protein and antibodies or antigen binding fragments of the disclosure in which the antibody or antigen binding fragment is fused to the C-terminus of the G protein or a biologically active portion thereof.

As used herein, a “targeted lipid particle” refers to a lipid particle that contains a targeted envelope protein embedded in the lipid bilayer, e.g., targeting CD8. Such targeted lipid particles can be a viral particle, a virus-like particle, a nanoparticle, a vesicle, an exosome, a dendrimer, a lentivirus, a viral vector, an enucleated cell, a

microvesicle, a membrane vesicle, an extracellular membrane vesicle, a plasma membrane vesicle, a giant plasma membrane vesicle, an apoptotic body, a mitoparticle, a pyrenocyte, a lysosome, another membrane enclosed vesicle, or a lentiviral vector, a viral based particle, a virus like particle (VLP), or a cell derived particle.

As used herein, a “retroviral nucleic acid” refers to a nucleic acid containing at least the minimal sequence requirements for packaging into a retrovirus or retroviral vector, alone or in combination with a helper cell, helper virus, or helper plasmid. In some embodiments, the retroviral nucleic acid further comprises or encodes an exogenous agent, a positive target cell-specific regulatory element, a non-target cell-specific regulatory element, or a negative TCSRE. In some embodiments, the retroviral nucleic acid comprises one or more of (e.g., all of) a 5' LTR (e.g., to promote integration), U3 (e.g., to activate viral genomic RNA transcription), R (e.g., a Tat-binding region), U5, a 3' LTR (e.g., to promote integration), a packaging site (e.g., ψ), and RRE (e.g., to bind to Rev and promote nuclear export). The retroviral nucleic acid can comprise RNA (e.g., when part of a virion) or DNA (e.g., when being introduced into a source cell or after reverse transcription in a recipient cell). In some embodiments, the retroviral nucleic acid is packaged using a helper cell, helper virus, or helper plasmid which comprises one or more of (e.g., all of) gag, pol, and env.

As used herein, a “target cell” refers to a cell of a type to which it is desired that a targeted lipid particle delivers an exogenous agent. In embodiments, a target cell is a cell of a specific tissue type or class, e.g., an immune effector cell, e.g., a T cell. In some embodiments, a target cell is a diseased cell, e.g., a cancer cell. In some embodiments, the fusogen, e.g., a re-targeted fusogen, leads to preferential delivery of the exogenous agent to a target cell compared to a non-target cell.

As used herein a “non-target cell” refers to a cell of a type to which it is not desired that a targeted lipid particle delivers an exogenous agent. In some embodiments, a non-target cell is a cell of a specific tissue type or class. In some embodiments, a non-target cell is a non-diseased cell, e.g., a non-cancerous cell. In some embodiments, the fusogen, e.g., a re-targeted fusogen, leads to lower delivery of the exogenous agent to a non-target cell compared to a target cell.

The term “effective amount” as used herein means an amount of a pharmaceutical composition which is sufficient to significantly and positively modify the symptoms and/or conditions to be treated (e.g., provide a positive clinical response). The effective amount of the targeted lipid particles of the disclosure for use in a pharmaceutical composition will vary with the particular condition being treated, the severity of the condition, the duration of treatment, the nature of concurrent therapy, the particular lipid particle) being employed, the particular pharmaceutically-acceptable excipient(s) and/or carrier(s) utilized, and like factors with the knowledge and expertise of the attending physician.

An “exogenous agent” as used herein with reference to a targeted lipid particle, refers to an agent that is neither comprised by nor encoded in the corresponding wild-type virus or fusogen made from a corresponding wild-type source cell. In some embodiments, the exogenous agent does not naturally exist, such as a protein or nucleic acid that has a sequence that is altered (e.g., by insertion, deletion, or substitution) relative to a naturally occurring protein. In some embodiments, the exogenous agent does not naturally exist in the source cell. In some embodiments, the exogenous agent exists naturally in the source cell but is exogenous to the virus. In some embodiments, the exogenous agent does not naturally exist in the recipient cell. In some embodiments, the exogenous agent exists naturally in the recipient cell, but is not present at a desired level or at a desired time. In some embodiments, the exogenous agent comprises DNA, RNA, or protein.

As used herein, a “promoter” refers to a cis-regulatory DNA sequence that, when operably linked to a gene coding sequence, drives transcription of the gene. The promoter may comprise one or more transcription factor binding sites. In some embodiments, a promoter works in concert with one or more enhancers which are distal to the gene.

As used herein, a composition refers to any mixture of two or more products, substances, or compounds, including cells. It may be a solution, a suspension, liquid, powder, a paste, aqueous, non-aqueous, or any combination thereof.

As used herein, the term “pharmaceutically acceptable” refers to a material, such as a carrier or diluent, which does not abrogate the biological activity or properties of a therapeutic compound, and is relatively nontoxic, i.e., the material may be administered to an individual without causing undesirable biological effects or

interacting in a deleterious manner with any of the components of the composition in which it is contained.

As used herein, the term “pharmaceutical composition” refers to a mixture of at least one targeted lipid particle of the disclosure with other chemical components, such as carriers, stabilizers, diluents, dispersing agents, suspending agents, thickening agents, and/or excipients. The pharmaceutical composition facilitates administration of the targeted lipid particle to an organism. Multiple techniques of administering targeted lipid particles of the disclosure exist in the art including, but not limited to, intravenous, oral, aerosol, parenteral, ophthalmic, pulmonary, and topical administration.

A “disease” or “disorder” as used herein refers to a condition in which treatment is needed and/or desired.

As used herein, the terms “treat,” “treating,” or “treatment” refer to ameliorating a disease or disorder, e.g., slowing or arresting or reducing the development of the disease or disorder or reducing at least one of the clinical symptoms thereof. For purposes of this disclosure, ameliorating a disease or disorder can include obtaining a beneficial or desired clinical result that includes, but is not limited to, any one or more of: alleviation of one or more symptoms, diminishment of extent of disease, preventing or delaying spread (for example, metastasis, for example metastasis to the lung or to the lymph node) of disease, preventing or delaying recurrence of disease, delay or slowing of disease progression, amelioration of the disease state, inhibiting the disease or progression of the disease, inhibiting or slowing the disease or its progression, arresting its development, and remission (whether partial or total). The terms “individual” and “subject” are used interchangeably herein to refer to an animal; for example a mammal. The terms include human and veterinary animals. In some embodiments, methods of treating animals, including, but not limited to, humans, rodents, simians, felines, canines, equines, bovines, porcines, ovines, caprines, mammalian laboratory animals, mammalian farm animals, mammalian sport animals, and mammalian pets, are provided. The animal can be male or female and can be any suitable age, including infant, juvenile, adolescent, adult, and geriatric. In some examples, an “individual” or “subject” refers to an animal in need of treatment for a disease or disorder. In some embodiments, the animal to receive the treatment can be a “patient,” designating the fact that the animal has been identified

as having a disorder of relevance to the treatment, or being at adequate risk of contracting the disorder. In particular embodiments, the animal is a human, such as a human patient.

CD8-Specific Antibodies

Described herein are novel antibodies and antigen binding fragments thereof that specifically target and bind CD8 α or CD8 β . In some embodiments, the antibodies or antigen binding fragments thereof may cross-react with cynomolgus (or “cyno”) or M. nemestrina CD8. In some embodiments, the antibodies or antigen binding fragments thereof are single-chain variable fragments (scFvs) composed of the antigen-binding domains derived from the heavy (VH) and the light (VL) chains of the IgG molecule and connected via a linker domain. In some embodiments, the antibodies or antigen binding fragments thereof are VHHs that correspond to the VH of the IgG molecule. The present disclosure also provides polynucleotides encoding the antibodies and fragments thereof, vectors, and host cells, and methods of using the antibodies or antigen binding fragments thereof. In some embodiments, e.g., the antibodies or antigen binding fragments thereof may be fused to henipavirus glycoprotein G for targeted binding and transduction to cells.

Sequences for exemplary antibodies and antigen binding fragments of the disclosure using the Kabat numbering scheme are shown in **Tables 3-4** below. Sequences for exemplary HCDRs of the disclosure are shown in **Table 3**. Sequences for exemplary LCDRs of the disclosure are shown in **Table 4**.

The sequences for the disclosed VH and VL domains are provided in **Tables 5-6**.

Tables 8-11 provided herein show the CDR sequences of the disclosed antibodies and antigen binding fragments thereof using both Chothia and IMGT numbering schemes. The full CD8 binder sequences of the variant CD8 scFvs and VHHs of the disclosure are shown in **Table 12**.

Table 3. HCDRS in Kabat Numbering Scheme

CD8 Binder	H-CDR1		H-CDR2		H-CDR3	
	Sequence	SEQ ID NO:	Sequence	SEQ ID NO:	Sequence	SEQ ID NO:
1	SYAIS	1	IIDPSDGNTNYAQNFQG	70	ERAAAGYYYYMDV	148
2	TYAIN	2	RIDPSSGGTKYAQNFQG	71	EHAAGTYYYYMDV	149
3	SYAIN	3	IIDPSGGNTNYAQNFQG	72	ERAAAGYYYYMDV	148
4	GYMH	4	HINPNNGDTNYAQNFQG	73	EGYYYYGMDV	649

5	DYIYQ	5	WINPNSGGTSYAQKFQG	74	EGDYYYGMDA	150
6	RYDIH	6	VINPNDGSTRYAQNFQG	75	ERGGMPDY	151
7	SYAMN	7	RINPNSGGTNYAQKFQG	76	GHGIPKY	152
8	SYIYH	8	WMNPNSGNTGYAQKFQG	77	VRSGSPQH	153
9	RHYIH	9	WMNPNSGNTGYAQKFQG	77	GGPWIVDAFDI	154
10	SYGIS	10	WISAHNGVTQYAQKFQG	78	GIAVAGTDY	650
11	NTDIN	11	IINPSGGSTSYAQKFQG	79	EATWGPYYYYMDV	155
12	RSYVH	12	WISPYNGNTKYAQKFQG	80	NKDGLQH	156
13	GYMH	13	IINPNSGDTKYAHQFQG	81	DAKRVGYYYYMDV	157
14	RYMH	14	RINPNSGGTNYAQKFQG	76	LVGGSPDY	158
15	NYDIN	15	RINPNSGGTNYAENFQG	1057	GAMVDY	159
16	NTDIN	11	IINPSDGDTKYAQEFQG	82	GNVGSYYYYGMDV	160
17	NYLH	16	WINPNSGDTKYAQKFQG	83	DSRGDWYFDL	161
18	RSIH	17	VIDPSGGSTSYAQKFQG	84	HGGRGLADY	162
19	SRDIS	18	WIDPKSGDTTYAQKFQG	85	LKELSSILDAFDI	163
20	SYDIN	19	MINPGAGSSTYAQKFQG	86	ERFGTGYYYYYMDV	164
21	NSDMN	20	LISGDGGTTYADSVKG	87	VIGEMVDDAFDL	165
22	GYMH	4	SINPNSGDTGYAQKFQG	88	ERLFGTYYYYYMDV	166
23	TYDIN	21	RIPIFGTANYAQKFQG	89	ADGELTDY	167
24	SYTMD	22	AIGTGGGIYADSVKG	90	HHLPAHYYYYYMDV	168
25	RYDIN	23	RINPNSGDTNYAQKFQG	91	DVPAGRYYYYYMDV	169
26	SYMH	24	MINPSDGSTRYAQKFQG	92	DRGVGRYYYYMDV	170
27	RYAVS	25	IINPSDGSTTYAQKFQG	93	DSRYGRYYYYMDV	171
28	NYAIS	26	IINPNGGSPSYAQKFQG	94	EIVVGPYYYYMDV	172
29	RYAIS	617	RINPNSGDTNYAQKFQG	91	GMVRGPYYYYMDV	173
30	SYAIS	1	IINPSGGSTSYAQTFQG	1058	EGVTGPYYYYMDV	174
31	RFDIN	28	IINPSDGSTDYAQNFQG	95	DAAAGTRYYYYYGMDV	175
32	SHAIS	29	IINPSGGSTSYAQKFQG	79	ELYSSTYYYYMDV	176
33	SYAIS	1	RINPNTGGTNHAQKFQG	96	ALYSGPYYYYYMDV	177
34	NSDMN	20	AISGSGGSTYADSVKG	97	EHAAGTYYYYYMDV	149
35	SYGIN	30	WISGYNGTDYARKLQG	98	DSLVGRYYYYYMDV	178
36	DYDIY	31	WISADNGNTNYEQKVQG	99	RSELDY	179
37	SYMH	32	WISPNSGATHYAQKFQG	100	GDDNDY	180
38	SYDIN	19	WINPNSGNTGYAKKFQG	101	GEEVDY	181
39	SYPMN	33	IINPSGGSTRYAQKFQG	102	GRRVPDY	182
40	DYIYH	34	WINPKSGATNYAQKFQG	103	GKVTTDY	183
41	SFEMN	35	RISESGDSSFYADSVKG	104	GRELIEY	184
42	DYAMH	36	AIGTGGGTYYADSVKG	105	VYDFPDV	185
43	DSYMH	37	WMNPNSGDTGYARKFQG	106	STYSHIDY	186
44	NYMH	38	TISPSDGSTTYAQRFGQ	107	EDSSGFDY	187
45	NYIYH	39	IINPSGGSTTYAQKFQG	108	DQGGGFDY	188
46	SYMH	40	GFDPEDGETIYAQKFQG	109	DQGWGMDV	189
47	SYIYH	8	RINPKSGRTYYAQNFGQ	110	LTEGIPDY	190
48	DYIYH	41	VINPGGGSTTYAQTFQG	111	DRYGPFDY	191
49	SYDIN	19	LMNPKTGDTNYAEKFQG	112	LVAGGAPDY	192
50	GYMH	4	IIDPSDGYTSYAQKFQG	113	DGFTGDIAY	193
51	GYMH	4	WINPNSGGTNYAQKFQG	114	VDDSSSPDY	194
52	GYYLH	42	GIMPISGTTIYAQKFQG	115	GPDGTEVDY	195
52	NHYMH	43	WMNPNSGNTGYAQKFQG	77	SESGDLDY	196
54	NYIYH	44	WMSPTSGDTGYAQKFQG	116	EVEIEGYMDV	197
55	SYMH	40	WINPNSGDTSYAQKFQG	117	DLDDDWYMDV	198
56	SYMH	40	IIDPSGDITSYAQKFQG	118	DSTTWDAFDI	199
57	DYMH	45	WINPNSGGTNYAQKFQG	114	VLVSGSPDY	200

58	ENEMH	46	IETSGGSTDYAQKFQG	119	EAAAGLDFQH	201
59	SYDMH	47	IINPNSGGTNYAQKLQG	120	ANSWDADY	202
60	NSDMH	48	VISGSGVTTYADSVKG	121	EHSSSWYTFDY	203
61	AYYMH	49	WINPNSGGTDYQKFQG	122	DDDSSGYLDY	204
62	NYIYH	44	MINPSGGSTTYAQKFQG	123	ASGDYMDLDYMDY	205
63	DYHMH	50	WINPDSGGTNYEQKFQG	124	VGSSGYLAPTH	206
64	DYYMH	51	WMNPNSGNTGYAQKFQG	77	VRGDGYNLGDY	207
65	DYYMH	52	WINPNSGGTNSAQKFQG	125	DVDTAMGAGDY	208
66	DYYIH	34	IINPSGGSASYAQKFQG	126	VARWGYGDYDPY	209
67	THDIN	53	IISPSDGSTSYAQKLQG	127	DRNGDYGGMDV	210
68	NYIYH	54	WINPISGGTHYQKFQG	128	EGLSSWYVLDY	211
69	SYDIN	19	WISADNGDTSFAQKFQG	129	DGSHYGYGMDV	212
70	SYDIN	19	GISPIYGTPYQKFQG	130	PGPEGYYGMDV	213
71	DNYMH	55	WMNPNSGNTGYAQKFQG	77	YHWDYGDYRFDY	214
72	SYIYH	8	WMNPNSGNTGYAQKFQG	77	VEIDYGDSPPDY	215
73	SYAIS	56	IINPSDGDTSYQKFQG	131	GAEWELRYAFDI	216
74	TYDIS	57	TINPSGGTTYAQKFQG	132	ETYYGLYYGMDV	217
75	SYDIN	19	WMNPKSGNTGYAQKFQG	133	APSLRGYSYGPDY	218
76	SYDIN	58	IINPSGGSTSYAQKFQG	79	DRQERYYYGMDV	219
77	SYDIN	19	IINPSDGGTDYQKFQG	1059	DRSYGDYYGMDV	220
78	SYDIN	58	IINPGGGNARHTQKFQG	134	EVFSENYYYGMDV	221
79	SYMH	40	IINPSDGGSTTYAQKFQG	93	EWDYTHYYGMDV	222
80	SHWIH	59	GFDPEDGETVYAQNFQG	135	GDSSGYQYFYDY	223
81	SYDIN	19	GITPVFGIANYAQKFQG	136	GSWDSSSWYIPEY	224
82	DYDIV	60	IINPRGGSTNYAQKFQG	137	LVWGGAYYYGMDV	225
83	SYGIS	10	WMNPNGDTDYQKFQG	138	PVFSGSYWYFDP	226
84	SYDIN	19	IINPSGGSTSYAQKFQG	139	DQAVAGPYYYGMDV	227
85	SYAIS	1	LINPGSGNTNYAQKFQG	140	DRWLAPYYGMDV	228
86	GHDMH	61	GIPIFGTPNYAQKFQG	141	VMGPVDYYGMDV	229
87	NYDMH	62	IINPSDGGSTTYAQKFQG	93	DLGPFGSYYGMDV	230
88	SYAMT	63	TINGDGDSTDYADSVKG	142	EGVVVPPYYGMDV	231
89	TYMH	64	QIDPNSGDTIYPQKFQG	143	SSGWSRYYGMDV	232
90	NYQIH	65	IINPSGGSTSYAQKFQG	79	DNGMTTGYGMDV	233
91	SYDIV	66	IINPSGGSTSYAQKFQG	79	DRAMVTGYGMDV	234
92	SYDIN	19	IVNPSDGNTNYAQKFQG	144	DRGYGDRGYGMDV	235
93	SYDIN	67	WINTYNGNTYQKLQG	145	SPKATADYYGMDV	236
94	SYDIN	19	IINPSDGITDYAQRFGQ	146	STVTPSYGMDV	237
95	SHAIH	68	IINPRDGDVYAQKFQG	147	EPVAGTGYGMDV	238
96	SYGIN	69	WMNPNSGNTGYAQKFQG	77	DNLAGFWSYGMDV	239
97	GYVMG	1061	AISRGGLSTYADSVKG	1062	DRSDLYEITAASNIDS	1063

Table 4. LCDRS in Kabat Numbering Scheme

CD8 Binder	L-CDR1		L-CDR2		L-CDR3	
	Sequence	SEQ ID NO:	Sequence	SEQ ID NO:	Sequence	SEQ ID NO:
1	RASQSISSYLN	240	AASSLQS	294	QQSYSTPLT	333
2	QASQDISNYLN	241	AASSLQS	294	QQSYSNLVS	334
3	RASQSISSYLN	240	AASSLQS	294	QQSYSTPLT	333
4	RSSQSLLSNGYNYLD	242	LGSNRAS	295	MQALQTPFT	335
5	RSSQSLLSNGYNYLD	242	LGSNRAS	295	MQGLQTPHT	336
6	RASQSISRNLN	243	KASNLKG	296	QQTYSAPL	337

7	RSSQSLHNSNGYNYLD	242	LGSNRAS	295	MQTLQTPLT	338
8	RASQSVASDLA	244	GASTRAT	297	QQYGDSPGS	339
9	QASQDIGNYLN	245	AATLQR	298	QQANSFPPT	340
10	RASQSISTHLA	246	GASTRAT	297	QQYGNRST	341
11	RASQTISNYLN	247	AATLQS	299	QQSYSTPPT	342
12	RASQGIRNDLG	248	DASTLQS	300	QQSYSSPYT	343
13	RASQISNYLN	249	AASSLQS	294	QQSYSTPYT	344
14	RSSQSLHNSNGYNYLD	242	LGSNRAS	295	MQGAHWPPPT	345
15	RASQGISDSLA	250	GASSLRS	301	QQSYRTPYT	346
16	RASQISNYLN	249	AASSLQS	294	QESFTTQWT	347
17	QASQDIHNYLN	251	DASNLET	302	QQANSFPPT	340
18	QASQDISNYLN	241	SASSLQS	303	QQRSNWPLYT	348
19	RASQISDWLA	252	AASSLQT	304	QQAISFPIT	349
20	QASQDISNYLN	241	SASTLQS	305	QQSYSSPFT	350
21	RASQSISTWLA	253	AATLQS	299	QQAISFPLT	351
22	RASQISNYLN	249	AATLQS	299	QQSYTFPIT	352
23	RSSQSLHNSNGYNYLD	242	DASHLET	306	QQYYSYPPT	353
24	QASQDISNYLN	241	AASTLHS	307	QQSYSAPLT	354
25	QASQDISNYLN	241	AATLQS	299	QQSFSTFYT	355
26	QASQDISNYLN	241	AATLQS	299	QQSYSIPFT	356
27	RASQSINRFLN	254	AASSLQN	308	QQSYSTPYT	344
28	RASQISSYLN	240	AASSLQS	294	QQSYSTPLT	333
29	QASQDISNYLN	241	AATLQS	299	QQSYSTPIT	357
30	RASQSVSTYLN	255	AASSLQS	294	QQSYTIPST	358
31	QASQDIAKYLN	256	AASSLQS	294	QQSYSAPPT	359
32	QASQGITNYLN	257	GASSLQS	309	QQSYSTPWT	360
33	RASQISSYLN	240	AASSLQS	294	QQSYSTPLT	333
34	QASQDIHNYLN	251	AATLQS	299	QQSYTTPLT	361
35	QASQDISNYLN	241	SAFSLQS	310	QQSYSAPIT	362
36	RASQISSYLN	240	SASNLQS	311	QQRSNWPPVT	363
37	QANQDISNLE	258	DASSLES	312	QQSYSIPIT	364
38	RASQGISNNLN	259	EASTLES	313	QQSYSTPLT	333
39	RSSQSLHNSNGYNYLD	242	GASTLET	314	MQGLQPPGT	365
40	RASQISRSLV	260	AATLQT	315	QQSYNHFRIT	366
41	QASQDISNYLN	241	DASNLET	302	QRSDSTPLT	367
42	QASHDISKSLN	261	GASTLQS	316	QQLNSYPRT	368
43	RASQDIGAYLA	262	AASSLQS	294	QQSYSIPYT	369
44	RASQISSYLA	263	AASSLQS	294	QQSYSTPYT	344
45	RASQGIRSYLA	264	GASNLET	317	QQSYSTPYT	344
46	RASQISSYLN	240	AASSLQS	294	QQTYSTPYT	370
47	RASQNIQTWLA	265	AATLQS	299	QQSYSTPQT	371
48	RASQTISYLN	266	AATLQS	299	QQSYRTPYT	346
49	RSSQSLHNSNGYNYLD	242	MGSNRAS	318	MQGTHWPT	372
50	RASQINNNYLN	267	GASSLQS	309	QQTFSLPYT	373
51	RASQTISTYLN	268	DASNLET	302	QQSYSTPYT	344
52	RASRGIGNDLA	269	DASTLET	319	QQGYNMPLT	374
52	RASQTIGNYVN	270	GASNLHT	320	QQTYSAPLT	375
54	RASQFIGSWLA	271	AATLQS	299	QQSYSPWT	376
55	RASQISSWMA	272	DASNLET	302	QQTYSTPYI	377
56	RASQGISNNLN	259	DASNLET	302	QQSYSSPWT	378
57	KSSQSVLYSSNNKNYLA	273	WASTRES	321	QQYASAPRT	379
58	RASQISSYLN	240	KTSSLES	322	QQSFTIPYT	380
59	RVSQGISSYLN	274	GASSLQS	309	QQSYSTPLT	333

60	RASQSIDWLA	252	DASNLET	302	QQSYSTPLT	333
61	RASQGISNYLA	275	SASNLQS	311	QQTYRTPPT	381
62	RASQSIRNYLT	276	SASNLQS	311	QQSYSTPLT	333
63	RASQNIRLYLN	277	AASTLQS	299	QQSLTTPFT	382
64	QASQDIRKFLN	278	AASSLQS	294	QQLNGYPGT	383
65	RASQSISSYLN	240	TASNLQS	323	QQSYSPLT	384
66	QASQDISNYLS	279	DASNLQS	324	QQTYTTPRT	385
67	RASQNVRSWLA	280	AASSLQS	294	QQSYNTPYT	386
68	RASQGIGNDLG	281	AASSLQS	294	QQSYAPPPT	387
69	RASQISNWLA	282	GASNLET	317	QQSYSTPPT	342
70	RSSQSLHSHNGYNYLD	242	LGSNRAS	295	MQGLQTPLT	388
71	RASQSISSYLN	240	LASSLQS	325	QQSDSIPVT	389
72	QASQDISNYLN	241	STSSLQS	326	QQSYSTPYN	390
73	RASEIGSWLA	283	AASSLQS	294	QQSYSTPYT	344
74	RASQISNYLN	249	AASSLQR	327	QQSYSTPLT	333
75	RASQSVTSNYLA	284	GASTRAT	297	QHYGSSPA	391
76	RASQSISSYLN	240	AASSLQS	294	QQSYSTPLT	333
77	RASQGISSYLA	285	AASTLQS	299	QQSYSTPPT	342
78	RASQDIGNYLN	286	AASSLQS	294	QQAYTYPYT	392
79	QASQDISNYLN	241	GASSLQS	309	QQSYTTPNT	393
80	RASQGISNYLA	275	AASTLQS	299	QQSYSTPYT	344
81	RASQGISNGLS	287	DASNLET	302	QQSYSTPFT	394
82	RASQNIRNYLN	288	GASSLQS	309	QQSYSTPLT	333
83	QASLDINNYLN	289	KASSLES	328	QQSYSMLT	395
84	QASQDISNYLN	241	AASSLQG	329	QQSYTTPWT	396
85	QASQDISNYLN	241	AASSLQS	294	QQSYSSPLT	397
86	QASQDISNYLN	241	KASSLES	328	QQSYSDPLT	398
87	QASQDISNYLN	241	GASTLQS	316	QQSYSAPIT	362
88	RASQISNYLN	249	AASNLS	330	QQSYTTPLT	361
89	RASQNIGNYLN	290	AASTLQS	299	QQSYSTPPWT	399
90	QASQDISNYLN	241	AASTLRS	331	QQSYQTPLT	400
91	QASQDISNYLN	241	AASTLQS	299	QQSYTTPPT	401
92	QASQDISNYLN	241	AASSLHS	332	QQSYSTPQT	371
93	RASQGIRNDLN	291	AASNLS	330	QQANSFPIT	402
94	RASQGINTWLA	292	AASSLQS	294	QQSYSTPYT	344
95	QASQDISNYLN	241	AASTLQS	299	QQSYTPPPT	403
96	RASQFIGSWLA	293	AASTLQS	299	QQDDSFPLT	404

In some embodiments, an antibody or antigen binding fragment thereof capable of binding CD8 α or CD8 β is disclosed, comprising a heavy chain variable region and a light chain variable region, wherein the heavy chain variable region comprises three heavy chain complementarity determining regions (HCDR1, HCDR2, and HCDR3), and the light chain variable region comprises three light chain complementarity determining regions (LCDR1, LCDR2, and LCDR3). In some embodiments, the HCDR1, HCDR2, and HCDR3 comprise amino acid sequences of any one of the SEQ ID NOs recited in **Tables 3, 8, and 10**, and the LCDR1, LCDR2, and LCDR3 comprise amino acid sequences of any one of the SEQ ID NOs recited in **Tables 3,**

9, and **11**. In some embodiments, the heavy chain variable region (VH) comprises an amino acid sequence of any one of SEQ ID NOs: 405-498 (**Table 5**) and the light chain variable region (VL) comprises an amino acid sequence of any one of SEQ ID NOs: 499-591 (**Table 6**).

In another embodiment, the antibody or antigen binding fragment thereof comprises a VH having an amino acid sequence with at least 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to a sequence selected from SEQ ID NOs: 405-498.

In another embodiment, the antibody or antigen binding fragment thereof comprises a VL having an amino acid sequence with at least 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to a sequence selected from SEQ ID NOs: 499-591.

In another embodiment, the antibody or antigen binding fragment comprises a VH having an amino acid sequence with at least 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to a sequence selected from SEQ ID NOs: 405-498 and a VL having an amino acid sequence with at least 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to a sequence selected from SEQ ID NOs: 499-591.

In another embodiment, the antibody or antigen binding fragment thereof comprises a VH having an amino acid sequence with at least 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 1060.

In another embodiment, the antibody or antigen binding fragment thereof comprises the VH of SEQ ID NO: 405 and the VL of SEQ ID NO: 499.

In another embodiment, the antibody or antigen binding fragment thereof comprises the VH of SEQ ID NO: 408 and the VL of SEQ ID NO: 503.

In another embodiment, the antibody or antigen binding fragment thereof comprises the VH of SEQ ID NO: 448 and the VL of SEQ ID NO: 542.

In another embodiment, the antibody or antigen binding fragment thereof comprises the VH of SEQ ID NO: 455 and the VL of SEQ ID NO: 549.

In another embodiment, the antibody or antigen binding fragment thereof comprises the VH of SEQ ID NO: 1060.

In another embodiment, the antibody or antigen binding fragment thereof comprises the HCDR1, HCDR2, HCDR3, LCDR1, LCDR2, and LCDR3 of SEQ ID NOs: 1, 70, 148, 240, 294, and 333, respectively.

In another embodiment, the antibody or antigen binding fragment thereof comprises the HCDR1, HCDR2, HCDR3, LCDR1, LCDR2, and LCDR3 of SEQ ID NOs: 5, 74, 150, 242, 295, and 336, respectively.

In another embodiment, the antibody or antigen binding fragment thereof comprises the HCDR1, HCDR2, HCDR3, LCDR1, LCDR2, and LCDR3 of SEQ ID NOs: 40, 109, 189, 240, 294, and 370, respectively.

In another embodiment, the antibody or antigen binding fragment thereof comprises the HCDR1, HCDR2, HCDR3, LCDR1, LCDR2, and LCDR3 of SEQ ID NOs: 43, 77, 196, 270, 320, and 375, respectively.

In another embodiment, the antibody or antigen binding fragment thereof comprises the HCDR1, HCDR2, and HCDR3 of SEQ ID NOs: 1061, 1062, and 1063, respectively.

In some embodiments, the single domain antibody can be human or humanized. In some embodiments, the single domain antibody or portion thereof is naturally occurring. In some embodiments, the single domain antibody or portion thereof is synthetic.

In some embodiments, the single domain antibodies are antibodies whose complementary determining regions are part of a single domain polypeptide. In some embodiments, the single domain antibody is a heavy chain only antibody variable domain. In some embodiments, the single domain antibody does not include light chains.

In various embodiments, any of the antibodies or antigen binding fragments described herein can comprise a heavy chain constant region and a light chain constant region. The heavy chain constant region may be an IgG, IgM, IgA, IgD, or IgE isotype, or a derivative or fragment thereof that retains at least one effector function of the intact heavy chain. The heavy chain constant region may be a human IgG isotype. The heavy chain constant region may be a human IgG1 or human IgG4 isotypes. The heavy chain constant region may be a human IgG1 isotype. The light chain constant region may be a human kappa light chain or lambda light chain or a derivative or fragment thereof that retains at least one effector function of the intact light chain. The light chain constant region may be a human kappa light chain.

In various embodiments, any of the disclosed antibodies or antigen binding fragments may be a rodent antibody or antigen binding fragment thereof, a chimeric

antibody or an antigen binding fragment thereof, a CDR-grafted antibody or an antigen binding fragment thereof, or a humanized antibody or an antigen binding fragment thereof. In another embodiment, any of the disclosed antibodies or antigen binding fragments comprises human or human-derived heavy and light chain variable regions, including human frameworks or human frameworks with one or more backmutations. In various embodiments, any of the disclosed antibodies or antigen binding fragments may be a Fab, Fab', F(ab')₂, Fd, scFv, (scFv)₂, scFv-Fc, VHH, or Fv fragment.

Antibodies whose heavy chain CDR, light chain CDR, VH, or VL amino acid sequences differ insubstantially from those shown in **Tables 3-6** are encompassed within the scope of the disclosure. Typically, this involves one or more conservative amino acid substitutions with an amino acid having similar charge, hydrophobic, or stereo chemical characteristics in the antigen-binding site or in the framework without adversely altering the properties of the antibody. Conservative substitutions may also be made to improve antibody properties, for example stability or affinity. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15 amino acid substitutions can be made to the VH or VL sequence. For example, a "conservative amino acid substitution" may involve a substitution of a native amino acid residue with a nonnative residue such that there is little or no effect on the polarity or charge of the amino acid residue at that position. Desired amino acid substitutions can be determined by those skilled in the art at the time such substitutions are desired. For example, amino acid substitutions can be used to identify important residues of the molecule sequence, or to increase or decrease the affinity of the molecules described herein. The following eight groups contain amino acids that are conservative amino acid substitutions for one another: 1) Alanine (A), Glycine (G); 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); 6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W); 7) Serine (S), Threonine (T); and 8) Cysteine (C), Methionine (M).

In some embodiments, the antibody or antigen binding fragment thereof binds to human CD8 α or CD8 β . In some embodiments, the antibody or antigen binding fragment thereof binds to a human CD8 α homodimer composed of two α chains. In some embodiments, the antibody or antigen binding fragment thereof binds to a human CD8 heterodimer composed of one α chain and one β chain.

In some embodiments, the antibody or antigen binding fragment binding CD8 is a single-chain variable fragment. In embodiments involving a single polypeptide containing both a heavy chain variable region and a light chain variable region, both orientations of these variable regions are contemplated. In some cases, the heavy chain variable region is on the N-terminal side of the light chain variable region, which means the heavy chain variable region is closer to the N-terminus of the polypeptide. In other cases, the light chain variable region is on the N-terminal side of the heavy chain variable region, which means the light chain variable region is closer to the N-terminus of the polypeptide than the heavy chain variable region.

In some embodiments, the scFv binding proteins comprise a linker. In some embodiments, the linker is between the heavy chain variable region (VH) and the light chain variable region (VL) (or vice versa). In some embodiments, the linker comprises the amino acid sequence of GS, GGS, GGS (SEQ ID NO: 645), GGGGS (SEQ ID NO: 627), GGGGS (SEQ ID NO: 625), any one of SEQ ID NOs: 645-648, or combinations thereof. Substitutions to introduce new disulfide bonds are also within the scope of the disclosure, e.g., by making substitutions G44C in the VH FR 2 and G100C in the VL FR4.

In some embodiments, the anti-CD8 antibody or antigen binding fragment binds to human CD8 with an affinity constant (K_D) of between about 1 nM and about 900 nM. In some embodiments, the K_D to human CD8 is between about 5 nM about 500 nM, about 6 nM to about 10 nM, about 11 nM to about 20 nM, about 25 nM to about 40 nM, about 40 nM to about 60 nM, about 70 nM to about 90 nM, about 100 nM to about 120 nM, about 125 nM to about 140 nM, about 145 nM to about 160 nM, about 170 nM and to about 200 nM, about 210 nM to about 250 nM, about 260 nM to about 300 nM, about 310 nM to about 350 nM, about 360 nM to about 400 nM, about 410 nM to about 450 nM, and about 460 nM to about 500 nM. In some embodiments, the anti-CD8 antibody or antigen binding fragment binds to human CD8 with an affinity constant (K_D) of 500 nM, 400 nM, 300 nM, 200 nM, 100 nM, 50 nM, 20 nM, or 10 nM or lower. In some embodiments, the anti-CD8 antibody or antigen binding fragment binds to human CD8 and cynomolgus, *M. mulatta* (rhesus monkey), or *M. nemestrina* CD8 with comparable binding affinity (K_D).

In some embodiments, the anti-CD8 antibody or antigen binding fragment binds to cynomolgus, *M. mulatta* (rhesus monkey), or *N. nemestrina* CD8. In some embodiments, the anti-CD8 antibody or antigen binding binds to mouse, dog, pig, etc., CD8. In some embodiments, the K_D to cynomolgus or *M. nemestrina* CD8 is between about 5 nM about 500 nM, about 6 nM to about 10 nM, about 11 nM to about 20 nM, about 25 nM to about 40 nM, about 40 nM to about 60 nM, about 70 nM to about 90 nM, about 100 nM to about 120 nM, about 125 nM to about 140 nM, about 145 nM to about 160 nM, about 170 nM and to about 200 nM, about 210 nM to about 250 nM, about 260 nM to about 300 nM, about 310 nM to about 350 nM, about 360 nM to about 400 nM, about 410 nM to about 450 nM, and about 460 nM to about 500 nM. In some embodiments, the anti-CD8 antibody or antigen binding fragment binds to cynomolgus or *M. nemestrina* CD8 with an affinity constant (K_D) of 500 nM, 400 nM, 300 nM, 200 nM, 100 nM, 50 nM, 20 nM, or 10 nM or lower.

An antibody or antigen binding fragment thereof that specifically binds CD8 α or CD8 β refers to an antibody or binding fragment that preferentially binds to CD8 α or CD8 β , respectively, over other antigen targets. As used herein, the term is interchangeable with an "anti-CD8" antibody or an "antibody that binds CD8." In some embodiments, the antibody or binding fragment capable of binding to CD8 α or CD8 β can do so with higher affinity for that antigen than others. In some embodiments, the antibody or binding fragment capable of binding CD8 α or CD8 β can bind to that antigen with a K_D of at least about 10^{-1} , 10^{-2} , 10^{-3} , 10^{-4} , 10^{-5} , 10^{-6} , 10^{-7} , 10^{-8} , 10^{-9} , 10^{-10} , 10^{-11} , 10^{-12} or greater (or any value in between), e.g., as measured by surface plasmon resonance or other methods known to those skilled in the art.

Another embodiment of the disclosure is an isolated polynucleotide encoding any of the antibody heavy chain variable regions or the antibody light chain variable regions of the disclosure. Certain exemplary polynucleotides are disclosed herein, however, other polynucleotides which, given the degeneracy of the genetic code or codon preferences in a given expression system, encode the antibodies or antigen binding fragments thereof of the disclosure are also within the scope of the disclosure. The polynucleotide sequences encoding a VH or a VL or a fragment thereof of the antibody or antigen binding fragments thereof of the disclosure can be operably linked to one or more regulatory elements, such as a promoter and enhancer, that

allow expression of the nucleotide sequence in the intended host cell. The polynucleotide may be a cDNA.

Another embodiment of the disclosure is a vector comprising the polynucleotide of the disclosure. Such vectors may be plasmid vectors, viral vectors, vectors for baculovirus expression, transposon-based vectors, or any other vector suitable for introduction of the polynucleotide of the disclosure into a given organism or genetic background by any means. For example, polynucleotides encoding light and heavy chain variable regions of the antibodies of the disclosure, optionally linked to constant regions, may be inserted into expression vectors. The light and heavy chains can be cloned in the same or different expression vectors. The DNA segments encoding immunoglobulin chains may be operably linked to control sequences in the expression vector(s) that ensure the expression of immunoglobulin polypeptides. Such control sequences include signal sequences, promoters (e.g., naturally associated or heterologous promoters), enhancer elements, and transcription termination sequences, and are chosen to be compatible with the host cell chosen to express the antibody. Once the vector has been incorporated into the appropriate host, the host is maintained under conditions suitable for high level expression of the proteins encoded by the incorporated polynucleotides.

Suitable expression vectors are typically replicable in the host organisms either as episomes or as an integral part of the host chromosomal DNA. Commonly, expression vectors contain selection markers such as ampicillin-resistance, hygromycin-resistance, tetracycline resistance, kanamycin resistance, or neomycin resistance to permit detection of those cells transformed with the desired DNA sequences. Suitable vectors, promoter, and enhancer elements are known in the art; many are commercially available for generating subject recombinant constructs.

Another embodiment of the disclosure is a host cell comprising the vector of the disclosure. The term "host cell" refers to a cell into which a vector has been introduced. It is understood that the term host cell is intended to refer not only to the particular subject cell but to the progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not be identical to the parent cell, but are still included within the scope of the term "host cell" as used herein. Such host cells may be eukaryotic cells, prokaryotic cells, plant cells, or archeal cells.

Escherichia coli, bacilli, such as *Bacillus subtilis*, and other enterobacteriaceae, such as *Salmonella*, *Serratia*, and various *Pseudomonas* species are examples of prokaryotic host cells. Other microbes, such as yeast, are also useful for expression. *Saccharomyces* (e.g., *S. cerevisiae*) and *Pichia* are examples of suitable yeast host cells. Exemplary eukaryotic cells may be of mammalian, insect, avian, or other animal origins.

Fusion Proteins Targeting CD8

Also provided herein are fusion proteins targeting CD8 that may be exposed on the surface on a lipid particle or viral vector. In some embodiments, the CD8 binders disclosed herein may be fused to an envelope glycoprotein G, H, and/or an F protein of the Paramyxoviridae family. In some embodiments the fusogen contains a Nipah virus protein F, a measles virus F protein, a tupaia paramyxovirus F protein, a paramyxovirus F protein, a Hendra virus F protein, a Henipavirus F protein, a Morbillivirus F protein, a respirovirus F protein, a Sendai virus F protein, a rubulavirus F protein, or an avulavirus F protein. In some embodiments, the lipid particle contains a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and/or a henipavirus envelope fusion glycoprotein F (F protein) or a biologically active portion thereof.

In particular embodiments, the fusogen is glycoprotein GP64 of baculovirus, or glycoprotein GP64 variant E45K/T259A.

In some embodiments, the fusogen is a hemagglutinin-neuraminidase (HN) and/or fusion (F) protein (F/HN) from a respiratory paramyxovirus. In some embodiments, the respiratory paramyxovirus is a Sendai virus. The HN and F glycoproteins of Sendai viruses function to attach to sialic acids via the HN protein, and to mediate cell fusion for entry into cells via the F protein. In some embodiments, the fusogen is a F and/or HN protein from the murine parainfluenza virus type 1 (See e.g., US Patent No. 10704061).

In some embodiments, the lipid particle (e.g. viral vector) is pseudotyped with viral glycoproteins as described herein such as a NiV-F and/or NiV-G protein.

In some embodiments, the viral vector further comprises a vector-surface targeting moiety which specifically binds to a target ligand. In some embodiments, the vector-surface targeting moiety is a polypeptide. In some embodiments, a nucleic acid

encoding the Paramyxovirus envelope protein (e.g. G protein) is modified with a targeting moiety to specifically bind to a target molecule on a target cells. In some embodiments, the targeting moiety can be any targeting protein, including but not necessarily limited to antibodies and antigen binding fragments thereof.

It has been reported that the henipavirus F proteins from various species exhibit compatibility with G proteins from other species to trigger fusion (Brandel-Tretheway et al. *Journal of Virology*. 2019. 93(13):e00577-19). In some aspects of the provided lipid particles (e.g. lentiviral vectors), the F protein is heterologous to the G protein, i.e. the F and G proteins or biologically active portions thereof are from different henipavirus species. For example, in some embodiments the G protein is from Hendra virus and the F protein is a NiV-F as described. In other aspects, the F and/or G protein are chimeric F and/or G protein containing regions of F and/or G proteins from different species of Henipavirus. In some embodiments, replacing a portion of the F protein with amino acids from a heterologous sequence of Henipavirus results in fusion to the G protein with the heterologous sequence. (Brandel-Tretheway et al. 2019). In some cases, the chimeric F and/or G protein contains an extracellular domain from one henipavirus species and a transmembrane and/or cytoplasmic domain from a different henipavirus species. For example, in some embodiments the F protein contains an extracellular domain of Hendra virus and a transmembrane/cytoplasmic domain of Nipah virus.

In some embodiments, the fusion protein contains a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and a single domain antibody (sdAb) variable domain or a single chain variable fragment (scFv). The sdAb variable domain or scFv can be linked directly or indirectly to the G protein. In particular embodiments, the sdAb variable domain or scFv is linked to the C-terminus (C-terminal amino acid) of the G protein or the biologically active portion thereof. The linkage can be via a peptide linker, such as a flexible peptide linker.

Table 7 provides a list of non-limiting examples of G proteins. Exemplary full length fusion protein sequences of the disclosure are disclosed in **Table 13**.

In some embodiments the G protein is a Henipavirus G protein or a biologically active portion thereof. In some embodiments, the Henipavirus G protein is a Hendra (HeV) virus G protein, a Nipah (NiV) virus G-protein (NiV-G), a Cedar (CedPV) virus G-protein, a Mojiang virus G-protein, a bat Paramyxovirus G-protein, or a biologically

active portion thereof. Non-limiting examples of G proteins include those corresponding to SEQ ID NOs: 609, 618, 619, 620, and 621.

In some embodiments, the attachment G proteins are type II transmembrane glycoproteins containing an N-terminal cytoplasmic tail (e.g., corresponding to amino acids 1-49 of SEQ ID NO: 600), a transmembrane domain (e.g., corresponding to amino acids 50-70 of SEQ ID NO: 600), and an extracellular domain containing an extracellular stalk (e.g., corresponding to amino acids 71-187 of SEQ ID NO: 600), and a globular head (corresponding to amino acids 188-602 of SEQ ID NO: 600). In such embodiments, the N-terminal cytoplasmic domain is within the inner lumen of the lipid bilayer and the C-terminal portion is the extracellular domain that is exposed on the outside of the lipid bilayer. Regions of the stalk in the C-terminal region (e.g. corresponding to amino acids 159-167 of NiV-G) have been shown to be involved in interactions with F protein and triggering of F protein fusion (Liu et al. 2015 J of Virology 89:1838). In wild-type G protein, the globular head mediates receptor binding to henipavirus entry receptors ephrin B2 and ephrin B3, but is dispensable for membrane fusion (Brandel-Tretheway et al. Journal of Virology. 2019. 93(13)e00577-19). In particular embodiments herein, tropism of the G protein is altered by linkage of the G protein or biologically active fragment thereof (e.g. cytoplasmic truncation) to a sdAb variable domain. Binding of the G protein to a binding partner can trigger fusion mediated by a compatible F protein or a biologically active portion thereof. G protein sequences disclosed herein are predominantly disclosed as expressed sequences including an N-terminal methionine required for start of translation. As such N-terminal methionines are commonly cleaved co- or post-translationally, the mature protein sequences for all G protein sequences disclosed herein are also contemplated as lacking the N-terminal methionine.

G glycoproteins are highly conserved among henipavirus species. For example, the G proteins of NiV and HeV viruses share 79% amino acid identity. Studies have shown a high degree of compatibility among G proteins with F proteins of different species as demonstrated by heterotypic fusion activation (Brandel-Tretheway et al. Journal of Virology. 2019). As described further below, a targeted lipid particle can contain heterologous G and F proteins from different species.

In some embodiments, the G protein has a sequence set forth in any of SEQ ID NOs: 600, 609, 618, 619, 620, 621, 628, 636, or 638-640, or is a functionally active variant or biologically active portion thereof that has a sequence that is at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% identical to any one of SEQ ID NOs: 600, 609, 618, 619, 620, 621, 628, 636, or 638-640. In particular embodiments, the G protein or functionally active variant or biologically active portion is a protein that retains fusogenic activity in conjunction with a Henipavirus F protein, such as an F protein (e.g. NiV-F or HeV-F). Fusogenic activity includes the activity of the G protein in conjunction with a Henipavirus F protein to promote or facilitate fusion of two membrane lumens, such as the lumen of the targeted lipid particle having embedded in its lipid bilayer a henipavirus F and G protein, and a cytoplasm of a target cell, e.g. a cell that contains a surface receptor or molecule that is recognized or bound by the targeted envelope protein. In some embodiments, the F protein and G protein are from the same Henipavirus species (e.g. NiV-G and NiV-F). In some embodiments, the F protein and G protein are from different Henipavirus species (e.g. NiV-G and HeV-F).

In particular embodiments, the G protein has the sequence of amino acids set forth in SEQ ID NOs: 600, 609, 618, 619, 620, 621, 628, 636, or 638-640, or is a functionally active variant thereof or a biologically active portion thereof that retains fusogenic activity. In some embodiments, the functionally active variant comprises an amino acid sequence having at least at or about 80%, at least at or about 85%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to any one of SEQ ID NOs: 600, 609, 618, 619, 620, 621, 628, 636, or 638-640 and retains fusogenic activity in conjunction with a Henipavirus F protein (e.g., NiV-F or HeV-F). In some embodiments, the biologically active portion has an amino acid sequence having at least at or about 80%, at least at or about

85%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to any one of SEQ ID NOs: 600, 609, 618, 619, 620, 621, 628, 636, or 638-640 and retains fusogenic activity in conjunction with a Henipavirus F protein (e.g., NiV-F or HeV-F).

Reference to retaining fusogenic activity includes activity (in conjunction with a Henipavirus F protein) that is at or about 10% to at or about 150% or more of the level or degree of binding of the corresponding wild-type G protein, such as set forth in any one of SEQ ID NOs: 600, 609, 618, 619, 620, 621, 628, 636, or 638-640, such as at least or at least about 10% of the level or degree of fusogenic activity of the corresponding wild-type G protein, such as at least or at least about 15% of the level or degree of fusogenic activity of the corresponding wild-type G protein, such as at least or at least about 20% of the level or degree of fusogenic activity of the corresponding wild-type G protein, such as at least or at least about 25% of the level or degree of fusogenic activity of the corresponding wild-type G protein, such as at least or at least about 30% of the level or degree of fusogenic activity of the corresponding wild-type G protein, such as at least or at least about 35% of the level or degree of fusogenic activity of the corresponding wild-type G protein, such as at least or at least about 40% of the level or degree of fusogenic activity of the corresponding wild-type G protein, such as at least or at least about 45% of the level or degree of fusogenic activity of the corresponding wild-type G protein, such as at least or at least about 50% of the level or degree of fusogenic activity of the corresponding wild-type G protein, such as at least or at least about 55% of the level or degree of fusogenic activity of the corresponding wild-type G protein, such as at least or at least about 60% of the level or degree of fusogenic activity of the corresponding wild-type G protein, such as at least or at least about 65% of the level or degree of fusogenic activity of the corresponding wild-type G protein, such as at least or at least about 70% of the level or degree of fusogenic activity of the corresponding wild-type G protein, such as at least or at least about 75% of the level or degree of fusogenic activity of the corresponding wild-type G protein, such as at least or at least about 80% of the level or degree of fusogenic activity of the corresponding wild-type G protein, such as at least or at least about 85% of the level

or degree of fusogenic activity of the corresponding wild-type G protein, such as at least or at least about 90% of the level or degree of fusogenic activity of the corresponding wild-type G protein, such as at least or at least about 95% of the level or degree of fusogenic activity of the corresponding wild-type G protein, such as at least or at least about 100% of the level or degree of fusogenic activity of the corresponding wild-type G protein, or such as at least or at least about 120% of the level or degree of fusogenic activity of the corresponding wild-type G protein.

In some embodiments, the G protein is a mutant G protein that is a functionally active variant or biologically active portion containing one or more amino acid mutations, such as one or more amino acid insertions, deletions, substitutions, or truncations. In some embodiments, the mutations described herein relate to amino acid insertions, deletions, substitutions, or truncations of amino acids compared to a reference G protein sequence. In some embodiments, the reference G protein sequence is the wild-type sequence of a G protein or a biologically active portion thereof. In some embodiments, the functionally active variant or the biologically active portion thereof is a mutant of a wild-type Hendra (HeV) virus G protein, a wild-type Nipah (NiV) virus G-protein (NiV-G), a wild-type Cedar (CedPV) virus G-protein, a wild-type Mojiang virus G-protein, a wild-type bat Paramyxovirus G-protein, or biologically active portions thereof. In some embodiments, the wild-type G protein has the sequence set forth in any one of SEQ ID NOs: 600, 609, 618, 619, 620, 621, 628, 636, or 638-640.

In some embodiments, the G protein is a mutant G protein that is a biologically active portion that is an N-terminally and/or C-terminally truncated fragment of a wild-type Hendra (HeV) virus G protein, a wild-type Nipah (NiV) virus G-protein (NiV-G), a wild-type Cedar (CedPV) virus G-protein, a wild-type Mojiang virus G-protein, or a wild-type bat Paramyxovirus G-protein. In particular embodiments, the truncation is an N-terminal truncation of all or a portion of the cytoplasmic domain. In some embodiments, the mutant G protein is a biologically active portion that is truncated and lacks up to 49 contiguous amino acid residues at or near the N-terminus of the wild-type G protein, such as a wild-type G protein set forth in any one of SEQ ID NOs: 600, 609, 618, 619, 620, 621, 628, 636, or 638-640. In some embodiments, the mutant G protein is truncated and lacks up to 49 contiguous amino acids, such as up to 49, 48, 47, 46, 45, 44, 43, 42, 41, 40, 30, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28,

27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2 or 1 contiguous amino acid(s) at the N-terminus of the wild-type G protein.

In some embodiments, the G protein is a wild-type Nipah virus G (NiV-G) protein or a Hendra virus G protein, or is a functionally active variant or biologically active portion thereof. In some embodiments, the G protein is a NiV-G protein that has the sequence set forth in SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628, or is a functional variant or a biologically active portion thereof that has an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628.

In some embodiments, the G protein is a mutant NiV-G protein that is a biologically active portion of a wild-type NiV-G. In some embodiments, the biologically active portion is an N-terminally truncated fragment. In some embodiments, the mutant NiV-G protein is truncated and lacks up to 5 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628), up to 6 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628), up to 7 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628), up to 8 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628), up to 9 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628), up to 10 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628), up to 11 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628), up to 12 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:600, SEQ ID NO:618, or SEQ

[illegible]

ID NO:600, SEQ ID NO:618, or SEQ ID NO:628), up to 31 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628), up to 32 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628), up to 33 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628), up to 34 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628), up to 35 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628), up to 36 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628), up to 37 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628), up to 38 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628), up to 39 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628), up to 40 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628), up to 41 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628), up to 42 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628), up to 43 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628), up to 44 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628), or up to 45 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628).

In some embodiments, the NiV-G protein is a biologically active portion that does not contain a cytoplasmic domain. In some embodiments, the NiV-G protein without the cytoplasmic domain is encoded by SEQ ID NO:622.

In some embodiments, the mutant NiV-G protein comprises a sequence set forth in any of SEQ ID NOs: 601-606, 629-634, 612, 622, or 637, or is a functional variant thereof that has an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NOs: 601-606, 629-634, 612, 622, or 637.

In some embodiments, the mutant NiV-G protein has a 5 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628), such as set forth in SEQ ID NO:601 or a functional variant thereof having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:601, or as set forth in SEQ ID NO:629 or a functional variant thereof having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:629 or a functional variant thereof having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at

least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:629.

In some embodiments, the mutant NiV-G protein has a 10 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628), such as set forth in SEQ ID NO:602 or a functional variant thereof having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:602, or such as set forth in SEQ ID NO:630 or a functional variant thereof having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:630.

In some embodiments, the mutant NiV-G protein has a 15 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628), such as set forth in SEQ ID NO:603 or a functional variant thereof that has an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:603, or such as set forth in SEQ ID NO:631 or a functional variant thereof having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at

or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:631.

In some embodiments, the mutant NiV-G protein has a 20 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628) such as set forth in SEQ ID NO:604, or a functional variant thereof having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:604, or such as set forth in SEQ ID NO:632 or a functional variant thereof having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:632.

In some embodiments, the mutant NiV-G protein has a 25 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628), such as set forth in SEQ ID NO:605 or a functional variant thereof having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:605, or such as set forth in

SEQ ID NO:633 or a functional variant thereof having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:633.

In some embodiments, the mutant NiV-G protein has a 30 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628), such as set forth in SEQ ID NO:606 or a functional variant thereof having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:606, or such as set forth in SEQ ID NO:634 or a functional variant thereof having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:634.

In some embodiments, the mutant NiV-G protein has a 33 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628) or a functional variant thereof having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about

93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:612, or such as set forth in SEQ ID NO:635 or a functional variant thereof having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:635.

In some embodiments, the mutant NiV-G protein has a 34 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628), such as set forth in SEQ ID NO:612 or a functional variant thereof having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:612, or such as set forth in SEQ ID NO:635 or a functional variant thereof having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:635.

In a preferred embodiment, the NiV-G protein has a 34 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO: 600, SEQ ID NO: 618, or SEQ ID NO: 628) and one or more amino acid substitutions corresponding to amino acid substitutions selected from E501A, W504A, Q530A, and E533A with reference to the numbering set forth in SEQ ID NO: 618.

In some embodiments, the mutant NiV-G protein lacks the N-terminal cytoplasmic domain of the wild-type NiV-G protein (SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628), such as set forth in SEQ ID NO:622 or a functional variant thereof having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:622.

In some embodiments, the mutant G protein is a mutant HeV-G protein that has the sequence set forth in SEQ ID NO:609 or 636, or is a functional variant or biologically active portion thereof that has an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:609 or 636.

In some embodiments, the G protein is a mutant HeV-G protein that is a biologically active portion of a wild-type HeV-G. In some embodiments, the biologically active portion is an N-terminally truncated fragment. In some embodiments, the mutant HeV-G protein is truncated and lacks up to 5 contiguous amino acid residues at or near the N-terminus of the wild-type HeV-G protein (SEQ ID NO:609 or 636), up to 6 contiguous amino acid residues at or near the N-terminus of the wild-type HeV-G protein (SEQ ID NO:609 or 636), up to 7 contiguous amino acid residues at or near the N-terminus of the wild-type HeV-G protein (SEQ ID NO:609 or 636), up to 8 contiguous amino acid residues at or near the N-terminus of the wild-type HeV-G protein (SEQ ID NO:609 or 636), up to 9 contiguous amino acid residues at or near the N-terminus of the wild-type HeV-G protein (SEQ ID NO:609 or 636), up to 10 contiguous amino acid residues at or near the N-terminus of the wild-type HeV-G protein (SEQ ID NO:609 or 636), up to 11 contiguous amino acid residues at or near

[illegible]

the N-terminus of the wild-type HeV-G protein (SEQ ID NO:609 or 636), up to 34 contiguous amino acid residues at or near the N-terminus of the wild-type HeV-G protein (SEQ ID NO:609 or 636), up to 35 contiguous amino acid residues at or near the N-terminus of the wild-type HeV-G protein (SEQ ID NO:609 or 636), up to 36 contiguous amino acid residues at or near the N-terminus of the wild-type HeV-G protein (SEQ ID NO:609 or 636), up to 37 contiguous amino acid residues at or near the N-terminus of the wild-type HeV-G protein (SEQ ID NO:609 or 636), up to 38 contiguous amino acid residues at or near the N-terminus of the wild-type HeV-G protein (SEQ ID NO:609 or 636), up to 39 contiguous amino acid residues at or near the N-terminus of the wild-type HeV-G protein (SEQ ID NO:609 or 636), up to 40 contiguous amino acid residues at or near the N-terminus of the wild-type HeV-G protein (SEQ ID NO:609 or 636), up to 41 contiguous amino acid residues at or near the N-terminus of the wild-type HeV-G protein (SEQ ID NO:609 or 636), up to 42 contiguous amino acid residues at or near the N-terminus of the wild-type HeV-G protein (SEQ ID NO:609 or 636), up to 43 contiguous amino acid residues at or near the N-terminus of the wild-type HeV-G protein (SEQ ID NO:609 or 636), up to 44 contiguous amino acid residues at or near the N-terminus of the wild-type HeV-G protein (SEQ ID NO:609 or 636), or up to 45 contiguous amino acid residues at or near the N-terminus of the wild-type HeV-G protein (SEQ ID NO:609 or 636).

In some embodiments, the HeV-G protein is a biologically active portion that does not contain a cytoplasmic domain. In some embodiments, the mutant HeV-G protein lacks the N-terminal cytoplasmic domain of the wild-type HeV-G protein (SEQ ID NO:609 or 636), such as set forth in SEQ ID NO:623 or a functional variant thereof having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:623.

In some embodiments, the G protein or the functionally active variant or biologically active portion thereof binds to Ephrin B2 or Ephrin B3. In some aspects, the G protein has the sequence of amino acids set forth in any one of SEQ ID NO:600,

SEQ ID NO:609, SEQ ID NO:618, SEQ ID NO:619, SEQ ID NO:628, SEQ ID NO:620, or SEQ ID NO:621, or is a functionally active variant thereof or a biologically active portion thereof that is able to bind to Ephrin B2 or Ephrin B3. In some embodiments, the functionally active variant or biologically active portion has an amino acid sequence having at least at or about 80%, at least at or about 85%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:600, SEQ ID NO:609, SEQ ID NO:618, SEQ ID NO:619, SEQ ID NO:628, SEQ ID NO:620, or SEQ ID NO:621, or a functionally active variant or biologically active portion thereof, and retains binding to Ephrin B2 or B3.

Reference to retaining binding to Ephrin B2 or B3 includes binding that is at least or at least about 5% of the level or degree of binding of the corresponding wild-type G protein, such as set forth in SEQ ID NO:600, SEQ ID NO:609, SEQ ID NO:618, SEQ ID NO:619, SEQ ID NO:628, SEQ ID NO:620, or SEQ ID NO:621, or a functionally active variant or biologically active portion thereof, 10% of the level or degree of binding of the corresponding wild-type G protein, such as set forth in SEQ ID NO:600, SEQ ID NO:609, SEQ ID NO:618, SEQ ID NO:619, SEQ ID NO:628, SEQ ID NO:620, or SEQ ID NO:621, or a functionally active variant or biologically active portion thereof, 15% of the level or degree of binding of the corresponding wild-type G protein, such as set forth in SEQ ID NO:600, SEQ ID NO:609, SEQ ID NO:618, SEQ ID NO:619, SEQ ID NO:628, SEQ ID NO:620, or SEQ ID NO:621, or a functionally active variant or biologically active portion thereof, 20% of the level or degree of binding of the corresponding wild-type G protein, such as set forth in SEQ ID NO:600, SEQ ID NO:609, SEQ ID NO:618, SEQ ID NO:619, SEQ ID NO:628, SEQ ID NO:620, or SEQ ID NO:621, or a functionally active variant or biologically active portion thereof, 25% of the level or degree of binding of the corresponding wild-type G protein, such as set forth in SEQ ID NO:600, SEQ ID NO:609, SEQ ID NO:618, SEQ ID NO:619, SEQ ID NO:628, SEQ ID NO:620, or SEQ ID NO:621, or a functionally active variant or biologically active portion thereof, 30% of the level or degree of binding of the corresponding wild-type G protein, such as set forth in SEQ ID NO:600, SEQ ID NO:609, SEQ ID NO:618, SEQ ID NO:619, SEQ ID NO:628, SEQ

ID NO:620, or SEQ ID NO:621, or a functionally active variant or biologically active portion thereof, 35% of the level or degree of binding of the corresponding wild-type G protein, such as set forth in SEQ ID NO:600, SEQ ID NO:609, SEQ ID NO:618, SEQ ID NO:619, SEQ ID NO:628, SEQ ID NO:620, or SEQ ID NO:621, or a functionally active variant or biologically active portion thereof, 40% of the level or degree of binding of the corresponding wild-type G protein, such as set forth in SEQ ID NO:600, SEQ ID NO:609, SEQ ID NO:618, SEQ ID NO:619, SEQ ID NO:628, SEQ ID NO:620, or SEQ ID NO:621, or a functionally active variant or biologically active portion thereof, 45% of the level or degree of binding of the corresponding wild-type G protein, such as set forth in SEQ ID NO:600, SEQ ID NO:609, SEQ ID NO:618, SEQ ID NO:619, SEQ ID NO:628, SEQ ID NO:620, or SEQ ID NO:621, or a functionally active variant or biologically active portion thereof, 50% of the level or degree of binding of the corresponding wild-type G protein, such as set forth in SEQ ID NO:600, SEQ ID NO:609, SEQ ID NO:618, SEQ ID NO:619, SEQ ID NO:628, SEQ ID NO:620, or SEQ ID NO:621, or a functionally active variant or biologically active portion thereof, 55% of the level or degree of binding of the corresponding wild-type G protein, such as set forth in SEQ ID NO:600, SEQ ID NO:609, SEQ ID NO:618, SEQ ID NO:619, SEQ ID NO:628, SEQ ID NO:620, or SEQ ID NO:621, or a functionally active variant or biologically active portion thereof, 60% of the level or degree of binding of the corresponding wild-type G protein, such as set forth in SEQ ID NO:600, SEQ ID NO:609, SEQ ID NO:618, SEQ ID NO:619, SEQ ID NO:628, SEQ ID NO:620, or SEQ ID NO:621, or a functionally active variant or biologically active portion thereof, 65% of the level or degree of binding of the corresponding wild-type G protein, such as set forth in SEQ ID NO:600, SEQ ID NO:609, SEQ ID NO:618, SEQ ID NO:619, SEQ ID NO:628, SEQ ID NO:620, or SEQ ID NO:621, or a functionally active variant or biologically active portion thereof, 70% of the level or degree of binding of the corresponding wild-type G protein, such as set forth in SEQ ID NO:600, SEQ ID NO:609, SEQ ID NO:618, SEQ ID NO:619, SEQ ID NO:628, SEQ ID NO:620, or SEQ ID NO:621, or a functionally active variant or biologically active portion thereof, such as at least or at least about 75% of the level or degree of binding of the corresponding wild-type G protein, such as set forth in SEQ ID NO:600, SEQ ID NO:609, SEQ ID NO:618, SEQ ID NO:619, SEQ ID NO:628, SEQ ID NO:620, or SEQ ID NO:621, or a functionally active variant or biologically active

portion thereof, such as at least or at least about 80% of the level or degree of binding of the corresponding wild-type G protein, such as set forth in SEQ ID NO:600, SEQ ID NO:609, SEQ ID NO:618, SEQ ID NO:619, SEQ ID NO:628, SEQ ID NO:620, or SEQ ID NO:621, or a functionally active variant or biologically active portion thereof, such as at least or at least about 85% of the level or degree of binding of the corresponding wild-type G protein, such as set forth in SEQ ID NO:600, SEQ ID NO:609, SEQ ID NO:618, SEQ ID NO:619, SEQ ID NO:628, SEQ ID NO:620, or SEQ ID NO:621, or a functionally active variant or biologically active portion thereof, such as at least or at least about 90% of the level or degree of binding of the corresponding wild-type G protein, such as set forth in SEQ ID NO:600, SEQ ID NO:609, SEQ ID NO:618, SEQ ID NO:619, SEQ ID NO:628, SEQ ID NO:620, or SEQ ID NO:621, or a functionally active variant or biologically active portion thereof, or such as at least or at least about 95% of the level or degree of binding of the corresponding wild-type protein, such as set forth in SEQ ID NO:600, SEQ ID NO:609, SEQ ID NO:618, SEQ ID NO:619, SEQ ID NO:628, SEQ ID NO:620, or SEQ ID NO:621, or a functionally active variant or biologically active portion thereof. In some embodiments, the G protein is NiV-G or a functionally active variant or biologically active portion thereof and binds to Ephrin B2 or Ephrin B3. In some aspects, the NiV-G has the sequence of amino acids set forth in SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628, or is a functionally active variant thereof or a biologically active portion thereof that is able to bind to Ephrin B2 or Ephrin B3. In some embodiments, the functionally active variant or biologically active portion has an amino acid sequence having at least at or about 80%, at least at or about 85%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628 and retains binding to Ephrin B2 or B3. Exemplary biologically active portions include N-terminally truncated variants lacking all or a portion of the cytoplasmic domain, e.g. 1 or more, such as 1 to 49 contiguous N-terminal amino acid residues, e.g. set forth in any one of SEQ ID NOS: 601-606, 622, and 629-634. Reference to retaining binding to Ephrin B2 or B3 includes binding that is at least or at least about 5% of the level or degree of binding of the corresponding wild-type

NiV-G, such as set forth in SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628, 10% of the level or degree of binding of the corresponding wild-type NiV-G, such as set forth in SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628, 15% of the level or degree of binding of the corresponding wild-type NiV-G, such as set forth in SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628, 20% of the level or degree of binding of the corresponding wild-type NiV-G, such as set forth in SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628, 25% of the level or degree of binding of the corresponding wild-type NiV-G, such as set forth in SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628, 30% of the level or degree of binding of the corresponding wild-type NiV-G, such as set forth in SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628, 35% of the level or degree of binding of the corresponding wild-type NiV-G, such as set forth in SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628, 40% of the level or degree of binding of the corresponding wild-type NiV-G, such as set forth in SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628, 45% of the level or degree of binding of the corresponding wild-type NiV-G, such as set forth in SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628, 50% of the level or degree of binding of the corresponding wild-type NiV-G, such as set forth in SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628, 55% of the level or degree of binding of the corresponding wild-type NiV-G, such as set forth in SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628, 60% of the level or degree of binding of the corresponding wild-type NiV-G, such as set forth in SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628, 65% of the level or degree of binding of the corresponding wild-type NiV-G, such as set forth in SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628, 70% of the level or degree of binding of the corresponding wild-type NiV-G, such as set forth in SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628, such as at least or at least about 75% of the level or degree of binding of the corresponding wild-type NiV-G, such as set forth in SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628, such as at least or at least about 80% of the level or degree of binding of the corresponding wild-type NiV-G, such as set forth in SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628, such as at least or at least about 85% of the level or degree of binding of the corresponding wild-type NiV-G, such as set forth in SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628, such

as at least or at least about 90% of the level or degree of binding of the corresponding wild-type NiV-G, such as set forth in SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628, or such as at least or at least about 95% of the level or degree of binding of the corresponding wild-type NiV-G, such as set forth in SEQ ID NO:600, SEQ ID NO:618, or SEQ ID NO:628.

In some embodiments, the G protein is HeV-G or a functionally active variant or biologically active portion thereof and binds to Ephrin B2 or Ephrin B3. In some aspects, the HeV-G has the sequence of amino acids set forth in SEQ ID NO:609 or 636, or is a functionally active variant thereof or a biologically active portion thereof that is able to bind to Ephrin B2 or Ephrin B3. In some embodiments, the functionally active variant or biologically active portion has an amino acid sequence having at least at or about 80%, at least at or about 85%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:609 or 636 and retains binding to Ephrin B2 or B3. Exemplary biologically active portions include N-terminally truncated variants lacking all or a portion of the cytoplasmic domain, e.g. 1 or more, such as 1 to 49 contiguous N-terminal amino acid residues, e.g. set forth in any one of SEQ ID NO:623.

Reference to retaining binding to Ephrin B2 or B3 includes binding that is at least or at least about 5% of the level or degree of binding of the corresponding wild-type HeV-G, such as set forth in SEQ ID NO:609 or 636, 10% of the level or degree of binding of the corresponding wild-type HeV-G, such as set forth in SEQ ID NO:609 or 636, 15% of the level or degree of binding of the corresponding wild-type HeV-G, such as set forth in SEQ ID NO:609 or 636, 20% of the level or degree of binding of the corresponding wild-type HeV-G, such as set forth in SEQ ID NO:609 or 636, 25% of the level or degree of binding of the corresponding wild-type HeV-G, such as set forth in SEQ ID NO:609 or 636, 30% of the level or degree of binding of the corresponding wild-type HeV-G, such as set forth in SEQ ID NO:609 or 636, 35% of the level or degree of binding of the corresponding wild-type HeV-G, such as set forth in SEQ ID NO:609 or 636, 40% of the level or degree of binding of the corresponding wild-type HeV-G, such as set forth in SEQ ID NO:609 or 636, 45% of the level or degree of binding of the corresponding wild-type HeV-G, such as set

forth in SEQ ID NO:609 or 636, 50% of the level or degree of binding of the corresponding wild-type HeV-G, such as set forth in SEQ ID NO:609 or 636, 55% of the level or degree of binding of the corresponding wild-type HeV-G, such as set forth in SEQ ID NO:609 or 636, 60% of the level or degree of binding of the corresponding wild-type HeV-G, such as set forth in SEQ ID NO:609 or 636, 65% of the level or degree of binding of the corresponding wild-type HeV-G, such as set forth in SEQ ID NO:609 or 636, 70% of the level or degree of binding of the corresponding wild-type HeV-G, such as set forth in SEQ ID NO:609 or 636, such as at least or at least about 75% of the level or degree of binding of the corresponding wild-type HeV-G, such as set forth in SEQ ID NO:609 or 636, such as at least or at least about 80% of the level or degree of binding of the corresponding wild-type NIV-G, such as set forth in SEQ ID NO:609 or 636, such as at least or at least about 85% of the level or degree of binding of the corresponding wild-type HeV-G, such as set forth in SEQ ID NO:609 or 636, such as at least or at least about 90% of the level or degree of binding of the corresponding wild-type HeV-G, such as set forth in SEQ ID NO:609 or 636, or such as at least or at least about 95% of the level or degree of binding of the corresponding wild-type HeV-G, such as set forth in SEQ ID NO:609 or 636.

In some embodiments, the G protein or the biologically thereof is a mutant G protein that exhibits reduced binding for the native binding partner of a wild-type G protein. In some embodiments, the mutant G protein or the biologically active portion thereof is a mutant of wild-type Niv-G and exhibits reduced binding to one or both of the native binding partners Ephrin B2 or Ephrin B3. In some embodiments, the mutant G-protein or the biologically active portion, such as a mutant NiV-G protein, exhibits reduced binding to the native binding partner. In some embodiments, the reduced binding to Ephrin B2 or Ephrin B3 is reduced by greater than at or about 5%, at or about 10%, at or about 15%, at or about 20%, at or about 25%, at or about 30%, at or about 40%, at or about 50%, at or about 60%, at or about 70%, at or about 80%, at or about 90%, or at or about 100%.

In some embodiments, the mutations described herein can improve transduction efficiency. In some embodiments, the mutations described herein allow for specific targeting of other desired cell types that are not Ephrin B2 or Ephrin B3. In some embodiments, the mutations described herein result in at least the partial inability to

bind at least one natural receptor, such as to reduce the binding to at least one of Ephrin B2 or Ephrin B3. In some embodiments, the mutations described herein interfere with natural receptor recognition.

In some embodiments, the mutant NiV-G protein or the biologically active portion thereof is truncated and lacks up to 5 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:618), 6 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:618), 7 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:618), 8 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:618), 9 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:618), 10 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:618), 11 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:618), 12 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:618), 13 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:618), 14 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:618), 15 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:618), 16 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:618), 17 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:618), 18 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:618), 19 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:618), 20 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:618), 21 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:618), 22 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:618), 23 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:618), 24 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:618), 25 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:618), 26 contiguous amino acid residues at or near the N-

terminus of the wild-type NiV-G protein (SEQ ID NO:618), 27 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:618), 28 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:618), 29 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:618), 30 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:618), 31 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:618), 32 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:618), 33 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:618), 34 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:618), 35 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:618), 36 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:618), 37 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:618), 38 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:618), 39 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:618), or 40 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:618).

In some embodiments, the G protein contains one or more amino acid substitutions in a residue that is involved in the interaction with one or both of Ephrin B2 and Ephrin B3. In some embodiments, the amino acid substitutions correspond to mutations E501A, W504A, Q530A, and E533A with reference to numbering set forth in SEQ ID NO:618.

In some embodiments, the G protein is a mutant G protein containing one or more amino acid substitutions selected from the group consisting of E501A, W504A, Q530A, and E533A with reference to numbering set forth in SEQ ID NO:618. In some embodiments, the G protein is a mutant G protein that contains one or more amino acid substitutions selected from the group consisting of E501A, W504A, Q530A, and E533A with reference to SEQ ID NO:618 or a biologically active portion thereof containing an N-terminal truncation. In some embodiments, the G protein is a mutant G protein that contains one or more amino acid substitutions selected from

the group consisting of E501A, W504A, Q530A, and E533A in combination with any one of the N-terminal truncations disclosed above with reference to SEQ ID NO:618 or a biologically active portion thereof. In some embodiments, any of the mutant G proteins described above contains one, two, three, or all four amino acid selected from the group consisting of E501A, W504A, Q530A, and E533A with reference to numbering set forth in SEQ ID NO:618, in all pairwise and triple combinations thereof.

In some embodiments, the mutant NiV-G protein has the amino acid sequence set forth in SEQ ID NO: 607 or 635 or an amino acid sequence having at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO: 607 or 635. In particular embodiments, the G protein has the sequence of amino acids set forth in SEQ ID NO: 607 or 635.

In some embodiments, the targeted envelope protein contains a G protein or a functionally active variant or biologically active portion thereof and an sdAb variable domain, in which the targeted envelope protein exhibits increased binding for another molecule that is different from the native binding partner of a wild-type G protein. In some embodiments, the other molecule can be a protein expressed on the surface of desired target cell. In some embodiments, the increased binding to the other molecule is increased by greater than at or about 25%, at or about 30%, at or about 40%, at or about 50%, at or about 60%, at or about 70%, at or about 80%, at or about 90%, or at or about 100%. In particular embodiments, the binding confers re-targeted binding compared to the binding of a wild-type G protein in which a new or different binding activity is conferred.

In some embodiments, the C-terminus of the single domain antibody is attached to the C-terminus of the G protein or biologically active portion thereof. In some embodiments, the N-terminus of the single domain antibody is exposed on the exterior surface of the lipid bilayer. In some embodiments, the N-terminus of the single domain antibody binds to a cell surface molecule of a target cell. In some embodiments, the single domain antibody specifically binds to a cell surface molecule present on a target cell. In some embodiments, the cell surface molecule is a protein, glycan, lipid, or low molecular weight molecule.

In some embodiments, the cell surface molecule of a target cell is an antigen or portion thereof. In some embodiments, the single domain antibody or portion thereof is an antibody having a single monomeric domain antigen binding/recognition domain that is able to bind selectively to a specific antigen. In some embodiments, the single domain antibody binds an antigen present on a target cell.

Exemplary cells include immune effector cells, peripheral blood mononuclear cells (PBMC) such as lymphocytes (T cells, B cells, natural killer cells) and monocytes, granulocytes (neutrophils, basophils, eosinophils), macrophages, dendritic cells, cytotoxic T lymphocytes, polymorphonuclear cells (also known as PMN, PML, or PMNL), stem cells, embryonic stem cells, neural stem cells, mesenchymal stem cells (MSCs), hematopoietic stem cells (HSCs), human myogenic stem cells, muscle-derived stem cells (MuStem), embryonic stem cells (ES or ESCs), limbal epithelial stem cells, cardio-myogenic stem cells, cardiomyocytes, progenitor cells, allogenic cells, resident cardiac cells, induced pluripotent stem cells (iPS), adipose-derived or phenotypic modified stem or progenitor cells, CD133+ cells, aldehyde dehydrogenase-positive cells (ALDH+), umbilical cord blood (UCB) cells, peripheral blood stem cells (PBSCs), neurons, neural progenitor cells, pancreatic beta cells, glial cells, or hepatocytes.

In some embodiments, the target cell is a cell of a target tissue. The target tissue can include liver, lungs, heart, spleen, pancreas, gastrointestinal tract, kidney, testes, ovaries, brain, reproductive organs, central nervous system, peripheral nervous system, skeletal muscle, endothelium, inner ear, or eye.

In some embodiments, the target cell is a muscle cell (e.g., skeletal muscle cell), kidney cell, liver cell (e.g. hepatocyte), or a cardiac cell (e.g. cardiomyocyte). In some embodiments, the target cell is a cardiac cell, e.g., a cardiomyocyte (e.g., a quiescent cardiomyocyte), a hepatoblast (e.g., a bile duct hepatoblast), an epithelial cell, a T cell (e.g. a naive T cell), a macrophage (e.g., a tumor infiltrating macrophage), or a fibroblast (e.g., a cardiac fibroblast).

In some embodiments, the target cell is a tumor-infiltrating lymphocyte, a T cell, a neoplastic or tumor cell, a virus-infected cell, a stem cell, a central nervous system (CNS) cell, a hematopoietic stem cell (HSC), a liver cell or a fully differentiated cell. In some embodiments, the target cell is a CD3+ T cell, a CD4+ T cell, a CD8+ T cell, a hepatocyte, a haematopoietic stem cell, a CD34+ haematopoietic stem cell, a

CD105+ haematopoietic stem cell, a CD117+ haematopoietic stem cell, a CD105+ endothelial cell, a B cell, a CD20+ B cell, a CD19+ B cell, a cancer cell, a CD133+ cancer cell, an EpCAM+ cancer cell, a CD19+ cancer cell, a Her2/Neu+ cancer cell, a GluA2+ neuron, a GluA4+ neuron, a NKG2D+ natural killer cell, a SLC1A3+ astrocyte, a SLC7A10+ adipocyte, or a CD30+ lung epithelial cell.

In some embodiments, the target cell is an antigen presenting cell, an MHC class II+ cell, a professional antigen presenting cell, an atypical antigen presenting cell, a macrophage, a dendritic cell, a myeloid dendritic cell, a plasmacyteoid dendritic cell, a CD11c+ cell, a CD11b+ cell, a splenocyte, a B cell, a hepatocyte, a endothelial cell, or a non-cancerous cell. In some embodiments, the cell surface molecule is any one of CD8.

In some embodiments, the G protein or functionally active variant or biologically active portion thereof is linked directly to the sdAb variable domain (e.g., a VHH) or scFv. In some embodiments, the targeted envelope protein is a fusion protein that has the following structure: (N'-single domain antibody-C')-(C'-G protein-N'). In some embodiments, the targeted envelope protein is a fusion protein that has the following structure: (N'-scFv-C')-(C'-G protein-N').

In some embodiments, the G protein or functionally active variant or biologically active portion thereof is linked indirectly via a linker to the the sdAb variable domain or scFv. In some embodiments, the linker is a peptide linker. In some embodiments, the linker is a chemical linker.

In some embodiments, the linker is a peptide linker and the targeted envelope protein is a fusion protein containing the G protein or functionally active variant or biologically active portion thereof linked via a peptide linker to the sdAb variable domain or svFv. In some embodiments, the targeted envelope protein is a fusion protein that has the following structure: (N'-single domain antibody-C')-Linker-(C'-G protein-N'). In some embodiments, the targeted envelope protein is a fusion protein that has the following structure: (N'-scFv-C')-Linker-(C'-G protein-N'). In some embodiments, the peptide linker is up to 65 amino acids in length. In some embodiments, the peptide linker comprises from or from about 2 to 65 amino acids, 2 to 60 amino acids, 2 to 56 amino acids, 2 to 52 amino acids, 2 to 48 amino acids, 2 to 44 amino acids, 2 to 40 amino acids, 2 to 36 amino acids, 2 to 32 amino acids, 2 to 28 amino acids, 2 to 24 amino acids, 2 to 20 amino acids, 2 to 18 amino acids, 2

to 14 amino acids, 2 to 12 amino acids, 2 to 10 amino acids, 2 to 8 amino acids, 2 to 6 amino acids, 6 to 65 amino acids, 6 to 60 amino acids, 6 to 56 amino acids, 6 to 52 amino acids, 6 to 48 amino acids, 6 to 44 amino acids, 6 to 40 amino acids, 6 to 36 amino acids, 6 to 32 amino acids, 6 to 28 amino acids, 6 to 24 amino acids, 6 to 20 amino acids, 6 to 18 amino acids, 6 to 14 amino acids, 6 to 12 amino acids, 6 to 10 amino acids, 6 to 8 amino acids, 8 to 65 amino acids, 8 to 60 amino acids, 8 to 56 amino acids, 8 to 52 amino acids, 8 to 48 amino acids, 8 to 44 amino acids, 8 to 40 amino acids, 8 to 36 amino acids, 8 to 32 amino acids, 8 to 28 amino acids, 8 to 24 amino acids, 8 to 20 amino acids, 8 to 18 amino acids, 8 to 14 amino acids, 8 to 12 amino acids, 8 to 10 amino acids, 10 to 65 amino acids, 10 to 60 amino acids, 10 to 56 amino acids, 10 to 52 amino acids, 10 to 48 amino acids, 10 to 44 amino acids, 10 to 40 amino acids, 10 to 36 amino acids, 10 to 32 amino acids, 10 to 28 amino acids, 10 to 24 amino acids, 10 to 20 amino acids, 10 to 18 amino acids, 10 to 14 amino acids, 10 to 12 amino acids, 12 to 65 amino acids, 12 to 60 amino acids, 12 to 56 amino acids, 12 to 52 amino acids, 12 to 48 amino acids, 12 to 44 amino acids, 12 to 40 amino acids, 12 to 36 amino acids, 12 to 32 amino acids, 12 to 28 amino acids, 12 to 24 amino acids, 12 to 20 amino acids, 12 to 18 amino acids, 12 to 14 amino acids, 14 to 65 amino acids, 14 to 60 amino acids, 14 to 56 amino acids, 14 to 52 amino acids, 14 to 48 amino acids, 14 to 44 amino acids, 14 to 40 amino acids, 14 to 36 amino acids, 14 to 32 amino acids, 14 to 28 amino acids, 14 to 24 amino acids, 14 to 20 amino acids, 14 to 18 amino acids, 18 to 65 amino acids, 18 to 60 amino acids, 18 to 56 amino acids, 18 to 52 amino acids, 18 to 48 amino acids, 18 to 44 amino acids, 18 to 40 amino acids, 18 to 36 amino acids, 18 to 32 amino acids, 18 to 28 amino acids, 18 to 24 amino acids, 18 to 20 amino acids, 20 to 65 amino acids, 20 to 60 amino acids, 20 to 56 amino acids, 20 to 52 amino acids, 20 to 48 amino acids, 20 to 44 amino acids, 20 to 40 amino acids, 20 to 36 amino acids, 20 to 32 amino acids, 20 to 28 amino acids, 20 to 26 amino acids, 20 to 24 amino acids, 24 to 65 amino acids, 24 to 60 amino acids, 24 to 56 amino acids, 24 to 52 amino acids, 24 to 48 amino acids, 24 to 44 amino acids, 24 to 40 amino acids, 24 to 36 amino acids, 24 to 32 amino acids, 24 to 30 amino acids, 24 to 28 amino acids, 28 to 65 amino acids, 28 to 60 amino acids, 28 to 56 amino acids, 28 to 52 amino acids, 28 to 48 amino acids, 28 to 44 amino acids, 28 to 40 amino acids, 28 to 36 amino acids, 28 to 34 amino acids, 28 to 32 amino acids, 32 to 65 amino acids, 32 to 60

amino acids, 32 to 56 amino acids, 32 to 52 amino acids, 32 to 48 amino acids, 32 to 44 amino acids, 32 to 40 amino acids, 32 to 38 amino acids, 32 to 36 amino acids, 36 to 65 amino acids, 36 to 60 amino acids, 36 to 56 amino acids, 36 to 52 amino acids, 36 to 48 amino acids, 36 to 44 amino acids, 36 to 40 amino acids, 40 to 65 amino acids, 40 to 60 amino acids, 40 to 56 amino acids, 40 to 52 amino acids, 40 to 48 amino acids, 40 to 44 amino acids, 44 to 65 amino acids, 44 to 60 amino acids, 44 to 56 amino acids, 44 to 52 amino acids, 44 to 48 amino acids, 48 to 65 amino acids, 48 to 60 amino acids, 48 to 56 amino acids, 48 to 52 amino acids, 50 to 65 amino acids, 50 to 60 amino acids, 50 to 56 amino acids, 50 to 52 amino acids, 54 to 65 amino acids, 54 to 60 amino acids, 54 to 56 amino acids, 58 to 65 amino acids, 58 to 60 amino acids, or 60 to 65 amino acids. In some embodiments, the peptide linker is a polypeptide that is 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, or 65 amino acids in length.

In particular embodiments, the linker is a flexible peptide linker. In some such embodiments, the linker is 1-20 amino acids, such as 1-20 amino acids predominantly composed of glycine. In some embodiments, the linker is 1-20 amino acids, such as 1-20 amino acids predominantly composed of glycine and serine. In some embodiments, the linker is a flexible peptide linker containing amino acids Glycine and Serine, referred to as GS-linkers. In some embodiments, the peptide linker includes the sequences GS, GGS, GGGGS (SEQ ID NO:627), GGGGGS (SEQ ID NO:625) or combinations thereof. In some embodiments, the polypeptide linker has the sequence (GGS) n , wherein n is 1 to 10. In some embodiments, the polypeptide linker has the sequence (GGGGS) n , (SEQ ID NO:626) wherein n is 1 to 10. In some embodiments, the polypeptide linker has the sequence (GGGGGS) n (SEQ ID NO:27), wherein n is 1 to 6.

Also provided herein are polynucleotides comprising a nucleic acid sequence encoding a targeted envelope protein. In some embodiments, the polynucleotides comprise a nucleic acid sequence encoding a G protein or biologically active portion thereof. In some embodiments, the polynucleotides further comprise a nucleic acid sequence encoding a single domain antibody (sdAb) variable domain or scFv or biologically active portion thereof. The polynucleotides may include a sequence of

nucleotides encoding any of the targeted envelope proteins described above. The polynucleotide can be a synthetic nucleic acid. Also provided are expression vectors containing any of the provided polynucleotides.

In some of any embodiments, expression of natural or synthetic nucleic acids is typically achieved by operably linking a nucleic acid encoding the gene of interest to a promoter and incorporating the construct into an expression vector. In some embodiments, vectors can be suitable for replication and integration in eukaryotes. In some embodiments, cloning vectors contain transcription and translation terminators, initiation sequences, and promoters useful for expression of the desired nucleic acid sequence. In some of any embodiments, a plasmid comprises a promoter suitable for expression in a cell.

In some embodiments, the polynucleotides contain at least one promoter that is operatively linked to control expression of the targeted envelope protein containing the G protein and the single domain antibody (sdAb) variable domain or scFv. For expression of the targeted envelope protein, at least one module in each promoter functions to position the start site for RNA synthesis. The best-known example of this is the TATA box, but in some promoters lacking a TATA box, such as the promoter for the mammalian terminal deoxynucleotidyl transferase gene and the promoter for the SV40 genes, a discrete element overlying the start site itself helps to fix the place of initiation.

In some embodiments, additional promoter elements, e.g., enhancers, regulate the frequency of transcriptional initiation. In some embodiments, additional promoter elements are located in the region 30-110 bp upstream of the start site, although a number of promoters have been shown to contain functional elements downstream of the start site as well. In some embodiments, spacing between promoter elements frequently is flexible, so that promoter function is preserved when elements are inverted or moved relative to one another. In some embodiments, such as with the thymidine kinase (tk) promoter, the spacing between promoter elements can be increased to 50 bp apart before activity begins to decline. In some embodiments, depending on the promoter, individual elements can function either cooperatively or independently to activate transcription.

A promoter may be one naturally associated with a gene or polynucleotide sequence, as may be obtained by isolating the 5' non-coding sequences located

upstream of the coding segment and/or exon. Such a promoter can be referred to as "endogenous." Similarly, an enhancer may be one naturally associated with a polynucleotide sequence, located either downstream or upstream of that sequence. Alternatively, certain advantages will be gained by positioning the coding polynucleotide segment under the control of a recombinant or heterologous promoter, which refers to a promoter that is not normally associated with a polynucleotide sequence in its natural environment. A recombinant or heterologous enhancer refers also to an enhancer not normally associated with a polynucleotide sequence in its natural environment. Such promoters or enhancers may include promoters or enhancers of other genes, and promoters or enhancers isolated from any other prokaryotic, viral, or eukaryotic cell, and promoters or enhancers not "naturally occurring," i.e., containing different elements of different transcriptional regulatory regions, and/or mutations that alter expression. In addition to producing nucleic acid sequences of promoters and enhancers synthetically, sequences may be produced using recombinant cloning and/or nucleic acid amplification technology, including PCR, in connection with the compositions disclosed herein.

In some embodiments, a suitable promoter is the immediate early cytomegalovirus (CMV) promoter sequence. In some embodiments, the promoter sequence is a strong constitutive promoter sequence capable of driving high levels of expression of any polynucleotide sequence operatively linked thereto. In some embodiments, a suitable promoter is Elongation Growth Factor-1 α (EF-1 α). In some embodiments, other constitutive promoter sequences may also be used, including, but not limited to the simian virus 40 (SV40) early promoter, mouse mammary tumor virus (MMTV), human immunodeficiency virus (HIV) long terminal repeat (LTR) promoter, MoMuLV promoter, an avian leukemia virus promoter, an Epstein-Barr virus immediate early promoter, a Rous sarcoma virus promoter, as well as human gene promoters such as, but not limited to, the actin promoter, the myosin promoter, the hemoglobin promoter, and the creatine kinase promoter.

In some embodiments, the promoter is an inducible promoter. In some embodiments, the inducible promoter provides a molecular switch capable of turning on expression of the polynucleotide sequence to which it is operatively linked when such expression is desired, or turning off the expression when expression is not desired. In some embodiments, inducible promoters comprise a metallothioneine

promoter, a glucocorticoid promoter, a progesterone promoter, and a tetracycline promoter.

In some embodiments, exogenously controlled inducible promoters can be used to regulate expression of the G protein and single domain antibody (sdAb) variable domain or scFv. For example, radiation-inducible promoters, heat-inducible promoters, and/or drug-inducible promoters can be used to selectively drive transgene expression in, for example, targeted regions. In such embodiments, the location, duration, and level of transgene expression can be regulated by the administration of the exogenous source of induction.

In some embodiments, expression of the targeted envelope protein containing a G protein and single domain antibody (sdAb) variable domain or scFv is regulated using a drug-inducible promoter. For example, in some cases, the promoter, enhancer, or transactivator comprises a Lac operator sequence, a tetracycline operator sequence, a galactose operator sequence, a doxycycline operator sequence, a rapamycin operator sequence, a tamoxifen operator sequence, or a hormone-responsive operator sequence, or an analog thereof. In some instances, the inducible promoter comprises a tetracycline response element (TRE). In some embodiments, the inducible promoter comprises an estrogen response element (ERE), which can activate gene expression in the presence of tamoxifen. In some instances, a drug-inducible element, such as a TRE, can be combined with a selected promoter to enhance transcription in the presence of drug, such as doxycycline. In some embodiments, the drug-inducible promoter is a small molecule-inducible promoter.

Any of the provided polynucleotides can be modified to remove CpG motifs and/or to optimize codons for translation in a particular species, such as human, canine, feline, equine, ovine, bovine, etc. species. In some embodiments, the polynucleotides are optimized for human codon usage (i.e., human codon-optimized). In some embodiments, the polynucleotides are modified to remove CpG motifs. In other embodiments, the provided polynucleotides are modified to remove CpG motifs and are codon-optimized, such as human codon-optimized. Methods of codon optimization and CpG motif detection and modification are well-known. Typically, polynucleotide optimization enhances transgene expression, increases transgene stability and preserves the amino acid sequence of the encoded polypeptide.

In order to assess the expression of the targeted envelope protein, the expression vector to be introduced into a cell can also contain either a selectable marker gene or a reporter gene or both to facilitate identification and selection of expressing particles, e.g. viral particles. In other embodiments, the selectable marker may be carried on a separate piece of DNA and used in a co-transfection procedure. Both selectable markers and reporter genes may be flanked with appropriate regulatory sequences to enable expression in the host cells. Useful selectable markers are known in the art and include, for example, antibiotic-resistance genes, such as neo and the like.

Reporter genes are used for identifying potentially transfected cells and for evaluating the functionality of regulatory sequences. Reporter genes that encode for easily assayable proteins are well known in the art. In general, a reporter gene is a gene that is not present in or expressed by the recipient organism or tissue and that encodes a protein whose expression is manifested by some easily detectable property, e.g., enzymatic activity. Expression of the reporter gene is assayed at a suitable time after the DNA has been introduced into the recipient cells.

Suitable reporter genes may include genes encoding luciferase, beta-galactosidase, chloramphenicol acetyl transferase, secreted alkaline phosphatase, or the green fluorescent protein gene (see, e.g., Ui-Tei et al., 2000, FEBS Lett. 479:79-82).

Suitable expression systems are well known and may be prepared using well known techniques or obtained commercially. Internal deletion constructs may be generated using unique internal restriction sites or by partial digestion of non-unique restriction sites. Constructs may then be transfected into cells that display high levels of the desired polynucleotide and/or polypeptide expression. In general, the construct with the minimal 5' flanking region showing the highest level of expression of reporter gene is identified as the promoter. Such promoter regions may be linked to a reporter gene and used to evaluate agents for the ability to modulate promoter-driven transcription.

Lipid Particles Targeting CD8

Also provided herein are targeted lipid particles (e.g. targeting CD8), such as targeted viral vectors, that comprise a henipavirus F protein molecule or biologically active portion thereof, and a fusion protein comprising (i) a henipavirus envelope

attachment glycoprotein G (G protein) or a biologically active portion thereof, and (ii) a single domain antibody (sdAb) variable domain or scFv, wherein the single domain antibody variable domain or scFv is attached to the C-terminus of the G protein or the biologically active portion, wherein each is exposed on the outer surface of the targeted viral vector. In particular embodiments, the provided targeted lipid particles exhibit fusogenic activity, which is mediated by the targeted envelope protein that facilitates binding to a target cell and contains the G protein or biologically active portion thereof, and the F protein or biologically active portion thereof that is involved in facilitating the merger or fusion of the two lumens of the lipid particle and the target cell membranes. **Table 7** provides non-limiting examples of G and F proteins for use in the targeted lipid particles of the disclosure.

In some embodiments, the targeted lipid particle provided herein (e.g. targeted lentiviral vector) has increased or greater expression of the targeted envelope protein compared to a reference lipid particle (e.g. reference lentiviral vector) that incorporates a similar envelope protein but that is fused to an alternative targeting moiety other than a sdAb variable domain, such as a single chain variable fragment (scFv). In some embodiments, the targeted viral vectors are produced by pseudotyping of viral vectors (e.g. lentiviral particles) following co-transfection of the packaging cells with the transfer, envelope, and gag-pol plasmids.

In some embodiments, the expression is increased by at or greater than 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 125%, 150%, 200%, 300%, 400%, 500% or more, compared to a reference lipid particle (e.g. reference lentiviral vector), e.g. a reference lipid particle containing a similar envelope protein but that is fused to an scFv. In some examples, the expression is increased by at or greater than 1.5-fold, 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold, 10-fold, 15-fold, 20-fold, 30-fold or more, compared to a reference lipid particle (e.g. reference lentiviral vector), e.g. a reference lipid particle containing a similar envelope protein but that is fused to an scFv. In some embodiments, expression can be assayed in vitro using flow cytometry, e.g. FACs. In some embodiments, expression can be depicted as the number or density of targeted envelope protein on the surface of a targeted viral vector (e.g. targeted lentiviral vector). In some embodiments, expression can be depicted as the mean fluorescent intensity (MFI) of surface expression of the targeted envelope protein on the surface of a targeted viral vector

(e.g. targeted lentiviral vector). In some embodiments, expression can be depicted as the percent of lipid particle (e.g. lentiviral vectors) in a population that are surface positive for the targeted envelope protein.

In some embodiments, in a population of targeted lipid particles (e.g. targeted lentiviral vectors) greater than at or about 50% of the lipid particles are surface positive for the targeted envelope protein. For example, in a population of provided targeted viral vectors (e.g. targeted lentiviral vectors) greater than at or about 55%, greater than at or about 60%, greater than at or about 65%, greater than at or about 70%, or greater than at or about 75% of the viral vectors in the population are surface positive for the targeted envelope protein.

In some embodiments, titer of the targeted lipid particles following introduction into target cells, such as by transduction (e.g. transduced cells), is increased compared to titer into the same target cells of reference lipid particles (e.g. reference lentiviral vector) that incorporate a similar envelope protein but fused to an alternative targeting moiety other than a sdAb variable domain, such as a single chain variable fragment (scFv). Typically, the alternative targeting moiety recognizes or binds the same target molecule as the sdAb variable domain of the targeted envelope protein of the targeted lipid particles. In some embodiments, the titer is increased by at or greater than 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 125%, 150%, 200%, 300%, 400%, 500% or more, compared to titer of a reference lipid particle (e.g. reference lentiviral vector), e.g. a reference lipid particle containing a similar envelope protein but that is fused to an scFv. In some examples, the titer is increased by at or greater than 1.5-fold, 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold, 10-fold, 15-fold, 20-fold, 30-fold or more, compared to the titer of a reference lipid particle (e.g. reference lentiviral vector), e.g. a reference viral vector containing a similar envelope protein but that is fused to an scFv. In some embodiments, the titer of the targeted viral vectors in target cells (e.g. transduced cells) is greater than at or about 1×10^6 transduction units (TU)/mL. For example, the titer of the targeted viral vectors in target cells (e.g. transduced cells) is greater than at or about 2×10^6 TU/mL, greater than at or about 3×10^6 TU/mL, greater than at or about 4×10^6 TU/mL, greater than at or about 5×10^6 TU/mL, greater than at or about 6×10^6 TU/mL, greater than at or about 7×10^6 TU/mL, greater than at or

about 8×10^6 TU/mL, greater than at or about 9×10^6 TU/mL, or greater than at or about 1×10^7 TU/mL.

A. *F proteins*

In some embodiments, the targeted lipid particle comprises one or more fusogens, e.g. henipavirus F proteins. In some embodiments, the targeted lipid particle contains an exogenous or overexpressed fusogen. In some embodiments, the fusogen is disposed in the lipid bilayer. In some embodiments, the fusogen facilitates the fusion of the targeted particle's lipid bilayer to a membrane. In some embodiments, the membrane is a plasma cell membrane.

In some embodiments, fusogens comprise protein based, lipid based, and chemical based fusogens. In some embodiments, the targeted lipid particle comprises a first fusogen comprising a protein fusogen and a second fusogen comprising a lipid fusogen or chemical fusogen. In some embodiments, the fusogen binds a fusogen binding partner on a target cell surface.

In some embodiments, the fusogen comprises a protein with a hydrophobic fusion peptide domain. In some embodiments, the fusogen comprises a henipavirus F protein molecule or biologically active portion thereof. In some embodiments, the Henipavirus F protein is a Hendra (Hev) virus F protein, a Nipah (NiV) virus F-protein, a Cedar (CedPV) virus F protein, a Mojiang virus F protein, a bat Paramyxovirus F protein, or a biologically active portion thereof.

In some embodiments, the N-terminal hydrophobic fusion peptide domain of the F protein molecule or biologically active portion thereof is exposed on the outside of a lipid bilayer.

F proteins of henipaviruses are encoded as F_0 precursors containing a signal peptide (e.g. corresponding to amino acid residues 1-26 of SEQ ID NO: 592). Following cleavage of the signal peptide, the mature F_0 (e.g. SEQ ID NO: 593) is transported to the cell surface, then endocytosed and cleaved by cathepsin L (e.g. between amino acids 109-110 of SEQ ID NO: 592) into the mature fusogenic subunits F1 (e.g. corresponding to amino acids 110-546 of SEQ ID NO: 592; set forth in SEQ ID NO: 595) and F2 (e.g. corresponding to amino acid residues 27-109 of SEQ ID NO: 1; set forth in SEQ ID NO: 594). The F1 and F2 subunits are associated by a disulfide bond and recycled back to the cell surface. The F1 subunit contains the fusion

peptide domain located at the N terminus of the F1 subunit (e.g., corresponding to amino acids 110-129 of SEQ ID NO:592) where it is able to insert into a cell membrane to drive fusion. In some cases, fusion activity is blocked by association of the F protein with G protein, until G engages with a target molecule resulting in its disassociation from F and exposure of the fusion peptide to mediate membrane fusion.

Among different henipavirus species, the sequence and activity of the F protein is highly conserved. For examples, the F protein of NiV and HeV viruses share 89% amino acid sequence identity. Further, in some cases, the henipavirus F proteins exhibit compatibility with G proteins from other species to trigger fusion (Brandel-Tretheway et al. *Journal of Virology*. 2019. 93(13):e00577-19). In some aspects of the provided targeted lipid particle, the F protein is heterologous to the G protein, i.e., the F and G protein or biologically active portions thereof are from different henipavirus species. For example, the F protein is from Hendra virus and the G protein is from Nipah virus. In other aspects, the F protein can be a chimeric F protein containing regions of F proteins from different species of Henipavirus. In some embodiments, switching a region of amino acid residues of the F protein from one species of Henipavirus to another can result in fusion to the G protein of the species comprising the amino acid insertion. (Brandel-Tretheway et al. 2019). In some cases, the chimeric F protein contains an extracellular domain from one henipavirus species and a transmembrane and/or cytoplasmic domain from a different henipavirus species. For example, the F protein may contain an extracellular domain of Hendra virus and a transmembrane/cytoplasmic domain of Nipah virus. F protein sequences disclosed herein are predominantly disclosed as expressed sequences including an N-terminal signal sequence. Such N-terminal signal sequences are commonly cleaved co- or post-translationally, thus the mature protein sequences for all F protein sequences disclosed herein are also contemplated as lacking the N-terminal signal sequence.

In some embodiments, the F protein is encoded by a nucleotide sequence that encodes the sequence set forth by any one of SEQ ID NOs: 592, 593, 608, 614-616, or 641-644, or is a functionally active variant or a biologically active portion thereof that has a sequence that is at least at or about 80%, at least at or about 85%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at

or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% identical to any one of SEQ ID NOS: 592, 593, 608, 614-616, or 641-644. In particular embodiments, the F protein or the functionally active variant or biologically active portion thereof retains fusogenic activity in conjunction with a Henipavirus G protein, such as a G protein set forth herein. Fusogenic activity includes the activity of the F protein in conjunction with a Henipavirus G protein to promote or facilitate fusion of two membrane lumens, such as the lumen of the targeted lipid particle having embedded in its lipid bilayer a henipavirus F and G protein, and a cytoplasm of a target cell, e.g., a cell that contains a surface receptor or molecule that is recognized or bound by the targeted envelope protein. In some embodiments, the F protein and G protein are from the same Henipavirus species (e.g. NiV-G and NiV-F). In some embodiments, the F protein and G protein are from different Henipavirus species (e.g., NiV-G and HeV-F). In particular embodiments, the F protein of the functionally active variant or biologically active portion retains the cleavage site cleaved by cathepsin L (e.g., corresponding to the cleavage site between amino acids 109-110 of SEQ ID NO:592).

In particular embodiments, the F protein has the sequence of amino acids set forth in SEQ ID NO:592, SEQ ID NO:593, SEQ ID NO:608, SEQ ID NO:614, SEQ ID NO:615, SEQ ID NO:616, SEQ ID NO:641, SEQ ID NO:642, SEQ ID NO:643, or SEQ ID NO:644 or is a functionally active variant thereof or a biologically active portion thereof that retains fusogenic activity. In some embodiments, the functionally active variant comprises an amino acid sequence having at least at or about 80%, at least at or about 85%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:592, SEQ ID NO:593, SEQ ID NO:608, SEQ ID NO:614, SEQ ID NO:615, SEQ ID NO:616, SEQ ID NO:641, SEQ ID NO:642, SEQ ID NO:643, or SEQ ID NO:644 and retains fusogenic activity in conjunction with a Henipavirus G protein (e.g., NiV-G or HeV-G). In some embodiments, the biologically active portion has an amino acid sequence having at least at or about 80%, at least at or about 85%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about

94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:592, SEQ ID NO:593, SEQ ID NO:608, SEQ ID NO:614, SEQ ID NO:615, SEQ ID NO:616, SEQ ID NO:641, SEQ ID NO:642, SEQ ID NO:643, or SEQ ID NO:644 and retains fusogenic activity in conjunction with a Henipavirus G protein (e.g., NiV-G or HeV-G). Reference to retaining fusogenic activity includes activity (in conjunction with a Henipavirus G protein) that is at or about 10% to at or about 150% or more of the level or degree of binding of the corresponding wild-type F protein, such as set forth in SEQ ID NO:592, SEQ ID NO:593, SEQ ID NO:608, SEQ ID NO:614, SEQ ID NO:615, SEQ ID NO:616, SEQ ID NO:641, SEQ ID NO:642, SEQ ID NO:643, or SEQ ID NO:644, such as at least or at least about 10% of the level or degree of fusogenic activity of the corresponding wild-type F protein, such as at least or at least about 15% of the level or degree of fusogenic activity of the corresponding wild-type F protein, such as at least or at least about 20% of the level or degree of fusogenic activity of the corresponding wild-type F protein, such as at least or at least about 25% of the level or degree of fusogenic activity of the corresponding wild-type F protein, such as at least or at least about 30% of the level or degree of fusogenic activity of the corresponding wild-type F protein, such as at least or at least about 35% of the level or degree of fusogenic activity of the corresponding wild-type F protein, such as at least or at least about 40% of the level or degree of fusogenic activity of the corresponding wild-type F protein, such as at least or at least about 45% of the level or degree of fusogenic activity of the corresponding wild-type F protein, such as at least or at least about 50% of the level or degree of fusogenic activity of the corresponding wild-type F protein, such as at least or at least about 55% of the level or degree of fusogenic activity of the corresponding wild-type f protein, such as at least or at least about 60% of the level or degree of fusogenic activity of the corresponding wild-type F protein, such as at least or at least about 65% of the level or degree of fusogenic activity of the corresponding wild-type F protein, such as at least or at least about 70% of the level or degree of fusogenic activity of the corresponding wild-type F protein, such as at least or at least about 75% of the level or degree of fusogenic activity of the corresponding wild-type F protein, such as at least or at least about 80% of the level or degree of fusogenic activity of the corresponding wild-type F protein, such as at least or at

least about 85% of the level or degree of fusogenic activity of the corresponding wild-type F protein, such as at least or at least about 90% of the level or degree of fusogenic activity of the corresponding wild-type F protein, such as at least or at least about 95% of the level or degree of fusogenic activity of the corresponding wild-type F protein, such as at least or at least about 100% of the level or degree of fusogenic activity of the corresponding wild-type F protein, or such as at least or at least about 120% of the level or degree of fusogenic activity of the corresponding wild-type F protein.

In some embodiments, the F protein is a mutant F protein that is a functionally active fragment or a biologically active portion containing one or more amino acid mutations, such as one or more amino acid insertions, deletions, substitutions, or truncations. In some embodiments, the mutations described herein relate to amino acid insertions, deletions, substitutions, or truncations of amino acids compared to a reference F protein sequence. In some embodiments, the reference F protein sequence is the wild-type sequence of an F protein or a biologically active portion thereof. In some embodiments, the mutant F protein or the biologically active portion thereof is a mutant of a wild-type Hendra (Hev) virus F protein, a Nipah (NiV) virus F-protein, a Cedar (CedPV) virus F protein, a Mojiang virus F protein, or a bat Paramyxovirus F protein. In some embodiments, the wild-type F protein is encoded by a sequence of nucleotides that encodes any one of SEQ ID NO: 592, 593, 608, 614-616, or 641-644.

In some embodiments, the mutant F protein is a biologically active portion of a wild-type F protein that is an N-terminally and/or C-terminally truncated fragment. In some embodiments, the mutant F protein or the biologically active portion of a wild-type F protein thereof comprises one or more amino acid substitutions. In some embodiments, the mutations described herein can improve transduction efficiency. In some embodiments, the mutations described herein can increase fusogenic capacity. Exemplary mutations include any as described, see e.g. Khetawat and Broder 2010 *Virology Journal* 7:312; Witting et al. 2013 *Gene Therapy* 20:997-1005; published international; patent application No. WO/2013/148327.

In some embodiments, the mutant F protein is a biologically active portion that is truncated and lacks up to 20 contiguous amino acid residues at or near the C-terminus of the wild-type F protein, such as a wild-type F protein encoded by a

sequence of nucleotides encoding the F protein set forth in any one of SEQ ID NOS: 592, 593, 608, or 614-616. In some embodiments, the mutant F protein is truncated and lacks up to 19 contiguous amino acids, such as up to 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, or 1 contiguous amino acid(s) at the C-terminus of the wild-type F protein.

In some embodiments, the F protein or the functionally active variant or biologically active portion thereof comprises an F1 subunit or a fusogenic portion thereof. In some embodiments, the F1 subunit is a proteolytically cleaved portion of the F0 precursor. In some embodiments, the F0 precursor is inactive. In some embodiments, the cleavage of the F0 precursor forms a disulfide-linked F1+F2 heterodimer. In some embodiments, the cleavage exposes the fusion peptide and produces a mature F protein. In some embodiments, the cleavage occurs at or around a single basic residue. In some embodiments, the cleavage occurs at Arginine 109 of NiV-F protein. In some embodiments, cleavage occurs at Lysine 109 of the Hendra virus F protein.

In some embodiments, the F protein is a wild-type Nipah virus F (NiV-F) protein or is a functionally active variant or biologically active portion thereof. In some embodiments, the F₀ precursor is encoded by a sequence of nucleotides encoding the sequence set forth in SEQ ID NO:592. The encoding nucleic acid can encode a signal peptide sequence that has the sequence M₁V₁V₁I₁L₁D₁K₁R₁C₁Y₁ C₁N₁L₁L₁I₁L₁I₁L₁M₁I₁ S₁E₁C₁S₁V₁G (SEQ ID NO:624) or another signal peptide sequence. In some embodiments, the F protein has the sequence set forth in SEQ ID NO:593. In some examples, the F protein is cleaved into an F1 subunit comprising the sequence set forth in SEQ ID NO:595 and an F2 subunit comprising the sequence set forth in SEQ ID NO:594.

In some embodiments, the F protein is a NiV-F protein that is encoded by a sequence of nucleotides encoding the sequence set forth in SEQ ID NO:592, or is a functionally active variant or biologically active portion thereof that has an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about

95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:592. In some embodiments, the NiV-F-protein has the sequence of set forth in SEQ ID NO:593, or is a functionally active variant or a biologically active portion thereof that has an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:593. In particular embodiments, the F protein or the functionally active variant or biologically active portion thereof retains the cleavage site cleaved by cathepsin L (e.g., corresponding to the cleavage site between amino acids 109-110 of SEQ ID NO:592).

In some embodiments, the F protein or the functionally active variant or the biologically active portion thereof includes an F1 subunit that has the sequence set forth in SEQ ID NO:595, or an amino acid sequence having, at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89% at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:595.

In some embodiments, the F protein or the functionally active variant or biologically active portion thereof includes an F2 subunit that has the sequence set forth in SEQ ID NO:594, or an amino acid sequence having, at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89% at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at

least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:594.

In some embodiments, the F protein is a mutant NiV-F protein that is a biologically active portion thereof that is truncated and lacks up to 20 contiguous amino acid residues at or near the C-terminus of the wild-type NiV-F protein (e.g., set forth SEQ ID NO:593). In some embodiments, the mutant NiV-F protein comprises an amino acid sequence set forth in SEQ ID NO:596. In some embodiments, the mutant NiV-F protein has a sequence that has at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:596. In some embodiments, the mutant F protein contains an F1 protein that has the sequence set forth in SEQ ID NO:597. In some embodiments, the mutant F protein has a sequence that has at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:597.

In some embodiments, the F protein is a mutant NiV-F protein that is a biologically active portion thereof that comprises a 20 amino acid truncation at or near the C-terminus of the wild-type NiV-F protein (SEQ ID NO:593); and a point mutation on an N-linked glycosylation site. In some embodiments, the mutant NiV-F protein comprises an amino acid sequence set forth in SEQ ID NO:598. In some embodiments, the mutant NiV-F protein has a sequence that has at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:598.

In some embodiments, the F protein is a mutant NiV-F protein that is a biologically active portion thereof that comprises a 22 amino acid truncation at or near the C-terminus of the wild-type NiV-F protein (SEQ ID NO:593). In some embodiments, the NiV-F protein comprises an amino acid sequence set forth in SEQ ID NO:599. In some embodiments, the NiV-F protein has a sequence with at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least

at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:599. In some embodiments, the NiV-F protein comprises an amino acid sequence set forth in SEQ ID NO:1092. In some embodiments, the NiV-F protein has a sequence with at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:1092. In some embodiments, the NiV-F protein comprises an amino acid sequence set forth in SEQ ID NO:1093. In some embodiments, the NiV-F protein has a sequence with at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:1093. In particular embodiments, the variant F protein is a mutant NiV-F protein that has the sequence of amino acids set forth in SEQ ID NO:613. In some embodiments, the NiV-F protein has a sequence with at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:613.

B. Lipid Bilayer

In some embodiments, the targeted lipid particle includes a naturally derived bilayer of amphipathic lipids that encloses a lumen or cavity. In some embodiments, the targeted lipid particle comprises a lipid bilayer as the outermost surface. In some embodiments, the lipid bilayer encloses a lumen. In some embodiments, the lumen is aqueous. In some embodiments, the lumen is in contact with the hydrophilic head groups on the interior of the lipid bilayer. In some embodiments, the lumen is a cytosol. In some embodiments, the cytosol contains cellular components present in a source cell. In some embodiments, the cytosol does not contain cellular components present in a source cell. In some embodiments, the lumen is a cavity. In some embodiments, the cavity contains an aqueous environment. In some embodiments, the cavity does not contain an aqueous environment.

In some aspects, the lipid bilayer is derived from a source cell during a process to produce a lipid-containing particle. In some embodiments, the lipid bilayer includes membrane components of the cell from which the lipid bilayer is produced, e.g., phospholipids, membrane proteins, etc. In some embodiments, the lipid bilayer includes a cytosol that includes components found in the cell from which the lipid bilayer is produced, e.g., solutes, proteins, nucleic acids, etc., but not all of the components of a cell, e.g., it lacks a nucleus. In some embodiments, the lipid bilayer is considered to be exosome-like. The lipid particle may vary in size, and in some instances have a diameter ranging from 30 and 300 nm, such as from 30 and 150 nm, and including from 40 to 100 nm.

In some embodiments, the lipid bilayer is a viral envelope. In some embodiments, the viral envelope is obtained from a source cell. In some embodiments, the viral envelope is obtained by the viral capsid from the source cell plasma membrane. In some embodiments, the lipid bilayer is obtained from a membrane other than the plasma membrane of a host cell. In some embodiments, the viral envelope lipid bilayer is embedded with viral proteins, including viral glycoproteins.

In other aspects, the lipid bilayer includes synthetic lipid complex. In some embodiments, the synthetic lipid complex is a liposome. In some embodiments, the lipid particle is a vesicular structure characterized by a phospholipid bilayer membrane and an inner aqueous medium. In some embodiments, the lipid bilayer has multiple lipid layers separated by aqueous medium. In some embodiments, the lipid bilayer forms spontaneously when phospholipids are suspended in an excess of aqueous solution. In some examples, the lipid components undergo self-rearrangement before the formation of closed structures and entrap water and dissolved solutes between the lipid bilayers.

In some embodiments, a targeted envelope protein and fusogen, such as any described above including any that are exogenous or overexpressed relative to the source cell, is disposed in the lipid bilayer.

In some embodiments, the targeted lipid particle comprises several different types of lipids. In some embodiments, the lipids are amphipathic lipids. In some embodiments, the amphipathic lipids are phospholipids. In some embodiments, the phospholipids comprise phosphatidylcholine, phosphatidylethanolamine, phosphatidylinositol, and phosphatidylserine. In some embodiments, the lipids

comprise phospholipids such as phosphocholines and phosphoinositols. In some embodiments, the lipids comprise DMPC, DOPC, and DSPC.

In some embodiments, the bilayer may be comprised of one or more lipids of the same or different type. In some embodiments, the source cell comprises a cell selected from CHO cells, BHK cells, MDCK cells, C3H 10T1/2 cells, FLY cells, Psi-2 cells, BOSC 23 cells, PA317 cells, WEHI cells, COS cells, BSC 1 cells, BSC 40 cells, BMT 10 cells, VERO cells, W138 cells, MRC5 cells, A549 cells, HT1080 cells, 293 cells, 293T cells, B-50 cells, 3T3 cells, NIH3T3 cells, HepG2 cells, Saos-2 cells, Huh7 cells, HeLa cells, W163 cells, 211 cells, and 211A cells.

C. Exogenous Agent

In some embodiments, the targeted lipid particle further comprises an agent that is exogenous relative to the source cell (also referred to herein as a “cargo” or “payload”). In some embodiments, the exogenous agent is a small molecule, a protein, or a nucleic acid (e.g., a DNA, a chromosome (e.g. a human artificial chromosome), an RNA, e.g., an mRNA or miRNA). In some embodiments, the exogenous agent or cargo encodes a cytosolic protein. In some embodiments the exogenous agent or cargo comprises or encodes a membrane protein. In some embodiments, the exogenous agent or cargo comprises a therapeutic agent. In some embodiments, the therapeutic agent is chosen from one or more of a protein, e.g., an enzyme, a transmembrane protein, a receptor, an antibody; a nucleic acid, e.g., DNA, a chromosome (e.g. a human artificial chromosome), RNA, mRNA, siRNA, miRNA; or a small molecule.

In some embodiments, the exogenous agent is present in at least, or no more than, 10, 20, 50, 100, 200, 500, 1,000, 2,000, 5,000, 10,000, 20,000, 50,000, 100,000, 200,000, 500,000, 1,000,000, 5,000,000, 10,000,000, 50,000,000, 100,000,000, 500,000,000, or 1,000,000,000 copies. In some embodiments, the targeted lipid particle has an altered, e.g., increased or decreased level of one or more endogenous molecules, e.g., protein or nucleic acid (e.g., in some embodiments, endogenous relative to the source cell, and in some embodiments, endogenous relative to the target cell), e.g., due to treatment of the source cell, e.g., mammalian source cell with a siRNA or gene editing enzyme. In some embodiments, the endogenous molecule is present in at least, or no more than, 10, 20, 50, 100, 200,

500, 1,000, 2,000, 5,000, 10,000, 20,000, 50,000, 100,000, 200,000, 500,000, 1,000,000, 5,000,000, 10,000,000, 50,000,000, 100,000,000, 500,000,000, or 1,000,000,000 copies. In some embodiments, the endogenous molecule (e.g., an RNA or protein) is present at a concentration of at least 1, 2, 3, 4, 5, 10, 20, 50, 100, 500, 10^3 , 5.0×10^3 , 10^4 , 5.0×10^4 , 10^5 , 5.0×10^5 , 10^6 , 5.0×10^6 , 1.0×10^7 , 5.0×10^7 , or 1.0×10^8 , greater than its concentration in the source cell. In some embodiments, the endogenous molecule (e.g., an RNA or protein) is present at a concentration of at least 1, 2, 3, 4, 5, 10, 20, 50, 100, 500, 10^3 , 5.0×10^3 , 10^4 , 5.0×10^4 , 10^5 , 5.0×10^5 , 10^6 , 5.0×10^6 , 1.0×10^7 , 5.0×10^7 , or 1.0×10^8 less than its concentration in the source cell.

In some embodiments, the targeted lipid particle (e.g., targeted viral vector) delivers to a target cell at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98%, or 99% of the cargo (e.g., a therapeutic agent, e.g., an exogenous therapeutic agent) comprised by the targeted lipid particle. In some embodiments, the targeted lipid particle that fuses with the target cell(s) delivers to the target cell an average of at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98%, or 99% of the cargo (e.g., a therapeutic agent, e.g., an exogenous therapeutic agent) comprised by the targeted lipid particle that fuses with the target cell(s). In some embodiments, the targeted lipid particle composition delivers to a target tissue at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98%, or 99% of the cargo (e.g., a therapeutic agent, e.g., an exogenous therapeutic agent) comprised by the targeted lipid particle composition.

In some embodiments, the exogenous agent or cargo is not expressed naturally in the cell from which the targeted lipid particle is derived. In some embodiments, the exogenous agent or cargo is expressed naturally in the cell from which the viral vector is derived. In some embodiments, the exogenous agent or cargo is loaded into the targeted lipid particle via expression in the cell from which the viral vector is derived (e.g. expression from DNA or mRNA introduced via transfection, transduction, or electroporation). In some embodiments, the exogenous agent or cargo is expressed from DNA integrated into the genome or maintained episomally. In some embodiments, expression of the exogenous agent or cargo is constitutive. In some embodiments, expression of the exogenous agent or cargo is induced. In some embodiments, expression of the exogenous agent or cargo is

induced immediately prior to generating the targeted lipid particle. In some embodiments, expression of the exogenous agent or cargo is induced at the same time as expression of the fusogen.

In some embodiments, the exogenous agent or cargo is loaded into the viral vector via electroporation into the viral vector itself or into the cell from which the viral vector is derived. In some embodiments, the exogenous agent or cargo is loaded into the viral vector via transfection (e.g., of a DNA or mRNA encoding the cargo) into the viral vector itself or into the cell from which the viral vector is derived.

In some embodiments, the exogenous agent or cargo may include one or more nucleic acid sequences, one or more polypeptides, a combination of nucleic acid sequences and/or polypeptides, one or more organelles, and any combination thereof. In some embodiments, the exogenous agent or cargo may include one or more cellular components. In some embodiments, the exogenous agent or cargo includes one or more cytosolic and/or nuclear components.

In some embodiments, the exogenous agent or cargo includes a nucleic acid, e.g., DNA, nDNA (nuclear DNA), mtDNA (mitochondrial DNA), protein coding DNA, gene, operon, chromosome, genome, transposon, retrotransposon, viral genome, intron, exon, modified DNA, mRNA (messenger RNA), tRNA (transfer RNA), modified RNA, microRNA, siRNA (small interfering RNA), tmRNA (transfer messenger RNA), rRNA (ribosomal RNA), mtRNA (mitochondrial RNA), snRNA (small nuclear RNA), small nucleolar RNA (snoRNA), SmY RNA (mRNA trans-splicing RNA), gRNA (guide RNA), TERC (telomerase RNA component), aRNA (antisense RNA), cis-NAT (Cis-natural antisense transcript), CRISPR RNA (crRNA), lncRNA (long noncoding RNA), piRNA (piwi-interacting RNA), shRNA (short hairpin RNA), tasiRNA (trans-acting siRNA), eRNA (enhancer RNA), satellite RNA, pcRNA (protein coding RNA), dsRNA (double stranded RNA), RNAi (interfering RNA), circRNA (circular RNA), reprogramming RNAs, aptamers, and any combination thereof. In some embodiments, the nucleic acid is a wild-type nucleic acid. In some embodiments, the nucleic acid is a mutant nucleic acid. In some embodiments the nucleic acid is a fusion or chimera of multiple nucleic acid sequences.

In some embodiments, the exogenous agent or cargo may include a nucleic acid. For example, the exogenous agent or cargo may comprise RNA to enhance expression of an endogenous protein, or a siRNA or miRNA that inhibits protein

expression of an endogenous protein. For example, the endogenous protein may modulate structure or function in the target cells. In some embodiments, the cargo may include a nucleic acid encoding an engineered protein that modulates structure or function in the target cells. In some embodiments, the exogenous agent or cargo is a nucleic acid that targets a transcriptional activator that modulate structure or function in the target cells.

In some embodiments, the exogenous agent or cargo includes a polypeptide, e.g., enzymes, structural polypeptides, signaling polypeptides, regulatory polypeptides, transport polypeptides, sensory polypeptides, motor polypeptides, defense polypeptides, storage polypeptides, transcription factors, antibodies, cytokines, hormones, catabolic polypeptides, anabolic polypeptides, proteolytic polypeptides, metabolic polypeptides, kinases, transferases, hydrolases, lyases, isomerases, ligases, enzyme modulator polypeptides, protein binding polypeptides, lipid binding polypeptides, membrane fusion polypeptides, cell differentiation polypeptides, epigenetic polypeptides, cell death polypeptides, nuclear transport polypeptides, nucleic acid binding polypeptides, reprogramming polypeptides, DNA editing polypeptides, DNA repair polypeptides, DNA recombination polypeptides, transposase polypeptides, DNA integration polypeptides, targeted endonucleases (e.g. Zinc -finger nucleases, transcription-activator-like nucleases (TALENs), cas9 and homologs thereof), recombinases, and any combination thereof. In some embodiments the protein targets a protein in the cell for degradation. In some embodiments the protein targets a protein in the cell for degradation by localizing the protein to the proteasome. In some embodiments, the protein is a wild-type protein. In some embodiments, the protein is a mutant protein. In some embodiments the protein is a fusion or chimeric protein.

In some embodiments, the exogenous agent or cargo includes a small molecule, e.g., ions (e.g. Ca^{2+} , Cl^- , Fe^{2+}), carbohydrates, lipids, reactive oxygen species, reactive nitrogen species, isoprenoids, signaling molecules, heme, polypeptide cofactors, electron accepting compounds, electron donating compounds, metabolites, ligands, and any combination thereof. In some embodiments the small molecule is a pharmaceutical that interacts with a target in the cell. In some embodiments the small molecule targets a protein in the cell for degradation. In some embodiments the small molecule targets a protein in the cell for degradation

by localizing the protein to the proteasome. In some embodiments that small molecule is a proteolysis targeting chimera molecule (PROTAC).

In some embodiments, the exogenous agent or cargo includes a mixture of proteins, nucleic acids, or metabolites, e.g., multiple polypeptides, multiple nucleic acids, multiple small molecules; combinations of nucleic acids, polypeptides, and small molecules; ribonucleoprotein complexes (e.g. Cas9-gRNA complex); multiple transcription factors, multiple epigenetic factors, reprogramming factors (e.g. Oct4, Sox2, cMyc, and Klf4); multiple regulatory RNAs; and any combination thereof.

In some embodiments, the exogenous agent or cargo includes one or more organelles, e.g., chondrisomes, mitochondria, lysosomes, nucleus, cell membrane, cytoplasm, endoplasmic reticulum, ribosomes, vacuoles, endosomes, spliceosomes, polymerases, capsids, acrosome, autophagosome, centriole, glycosome, glyoxysome, hydrogenosome, melanosome, mitosome, myofibril, cnidocyst, peroxisome, proteasome, vesicle, stress granule, networks of organelles, and any combination thereof.

In some embodiments, the exogenous agent encodes a therapeutic agent or a diagnostic agent. In some embodiments, the therapeutic agent is a chimeric antigen receptor (CAR) or T-cell receptor (TCR). In some embodiments, the CAR targets a tumor antigen selected from CD19, CD20, CD22, or BCMA. In another embodiment, the CAR is engineered to comprise an intracellular signaling domain of the T cell antigen receptor complex zeta chain (e.g., CD3 zeta). In a preferred embodiment, the intracellular domain is selected from a CD137 (4-1BB) signaling domain, a CD28 signaling domain, and a CD3zeta signaling domain.

D. Methods of Generating Targeted Lipid Particles Derived from Virus

Provided herein are targeted lipid particles that are derived from virus, such as viral particles or virus-like particles, including those derived from retroviruses or lentiviruses. In some embodiments, the targeted lipid particle's bilayer of amphipathic lipids is or comprises the viral envelope. In some embodiments, the targeted lipid particle's bilayer of amphipathic lipids is or comprises lipids derived from a producer cell. In some embodiments, the viral envelope may comprise a fusogen, e.g., a fusogen that is endogenous to the virus or a pseudotyped fusogen. In some embodiments, the targeted lipid particles's lumen or cavity comprises a viral nucleic

acid, e.g., a retroviral nucleic acid, e.g., a lentiviral nucleic acid. In some embodiments, the viral nucleic acid may be a viral genome. In some embodiments, the targeted lipid particle further comprises one or more viral non-structural proteins, e.g., in its cavity or lumen. In some embodiments, the targeted lipid particles is or comprises a virus-like particle (VLP). In some embodiments, the VLP does not comprise an envelope. In some embodiments, the VLP comprises an envelope. In some embodiments, the viral particle or virus-like particle, such as a retrovirus or retrovirus-like particle, comprises one or more of a Gag polypeptide, polymerase (e.g., Pol), integrase (IN, e.g., a functional or non-functional variant), protease (PR), and a fusogen. In some embodiments, the targeted lipid particle further comprises Rev. In some embodiments, one or more of the aforesaid proteins are encoded in the retroviral genome, and in some embodiments, one or more of the aforesaid proteins are provided in trans, e.g., by a helper cell, helper virus, or helper plasmid. In some embodiments, the targeted lipid particle nucleic acid (e.g., retroviral nucleic acid) comprises one or more of the following nucleic acid sequences: 5' LTR (e.g., comprising U5 and lacking a functional U3 domain), Psi packaging element (Psi), Central polypurine tract (cPPT) Promoter operatively linked to the payload gene, payload gene (optionally comprising an intron before the open reading frame), Poly A tail sequence, WPRE, and 3' LTR (e.g., comprising U5 and lacking a functional U3). In some embodiments the targeted lipid particle nucleic acid further comprises one or more insulator elements. In some embodiments, the recognition sites are situated between the poly A tail sequence and the WPRE.

In some embodiments, the targeted lipid particle comprises supramolecular complexes formed by viral proteins that self-assemble into capsids. In some embodiments, the targeted lipid particle is a viral particle or virus-like particle derived from viral capsids. In some embodiments, the targeted lipid particle is a viral particle or virus-like particle derived from viral nucleocapsids. In some embodiments, the targeted lipid particle comprises nucleocapsid-derived proteins that retain the property of packaging nucleic acids. In some embodiments, the viral particles or virus-like particles comprises only viral structural glycoproteins. In some embodiments, the targeted lipid particle does not contain a viral genome. In some embodiments, the targeted lipid particle packages nucleic acids from host cells during the expression process. In some embodiments, the nucleic acids do not

encode any genes involved in virus replication. In particular embodiments, the targeted lipid particle is a virus-like particle, e.g. retrovirus-like particle such as a lentivirus-like particle, that is replication defective.

In some cases, the targeted lipid particle is a viral particle that is morphologically indistinguishable from the wild type infectious virus. In some embodiments, the viral particle presents the entire viral proteome as an antigen. In some embodiments, the viral particle presents only a portion of the proteome as an antigen.

In some embodiments, the viral particle or virus-like particle is produced utilizing proteins (e.g., envelope proteins) from a virus within the Paramyxoviridae family. In some embodiments, the Paramyxoviridae family comprises members within the Henipavirus genus. In some embodiments, the Henipavirus is or comprises a Hendra (HeV) or a Nipah (NiV) virus. In particular embodiments, the viral particles or virus-like particles incorporate a targeted envelope protein and fusogen.

In some embodiments, viral particles or virus-like particles may be produced in multiple cell culture systems including bacteria, mammalian cell lines, insect cell lines, yeast, and plant cells.

Suitable cell lines which can be used include, for example, CHO cells, BHK cells, MDCK cells, C3H 10T1/2 cells, FLY cells, Psi-2 cells, BOSC 23 cells, PA317 cells, WEHI cells, COS cells, BSC 1 cells, BSC 40 cells, BMT 10 cells, VERO cells, W138 cells, MRC5 cells, A549 cells, HT1080 cells, 293 cells, 293T cells, B-50 cells, 3T3 cells, NIH3T3 cells, HepG2 cells, Saos-2 cells, Huh7 cells, HeLa cells, W163 cells, 211 cells, 211A cells, and cyno and Macaca nemestrina cell lines. In embodiments, the packaging cells are 293 cells, 293T cells, or A549 cells.

In some embodiments, a source cell line includes a cell line which is capable of producing recombinant retroviral particles, comprising a producer cell line and a transfer vector construct comprising a packaging signal. Methods of preparing viral stock solutions are illustrated by, e.g., Y. Soneoka et al. (1995) *Nucl. Acids Res.* 23:628-633, and N. R. Landau et al. (1992) *J. Virol.* 66:5110-5113, which are incorporated herein by reference.

In some embodiments, the assembly of a viral particle or virus-like particle is initiated by binding of the core protein to a unique encapsidation sequence within the viral genome (e.g. UTR with stem-loop structure). In some embodiments, the interaction of the core with the encapsidation sequence facilitates oligomerization.

In some embodiments, the targeted lipid particle is a virus-like particle which comprises a sequence that is devoid of or lacking viral RNA. In some embodiments, such particles may be the result of removing or eliminating the viral RNA from the sequence. In some embodiments, this may be achieved by using an endogenous packaging signal binding site on Gag. In some embodiments, the endogenous packaging signal binding site is on Pol. In some embodiments, the RNA which is to be delivered will contain a cognate packaging signal. In some embodiments, a heterologous binding domain (which is heterologous to Gag) located on the RNA to be delivered, and a cognate binding site located on Gag or Pol, can be used to ensure packaging of the RNA to be delivered. In some embodiments, the heterologous sequence could be non-viral or it could be viral, in which case it may be derived from the same virus or a different virus. In some embodiments, the vector particles could be used to deliver therapeutic RNA, in which case functional integrase and/or reverse transcriptase is not required. In some embodiments, the vector particles could also be used to deliver a therapeutic gene of interest, in which case Pol is typically included. In some embodiments, the retroviral nucleic acid comprises one or more of (e.g., all of): a 5' promoter (e.g., to control expression of the entire packaged RNA), a 5' LTR (e.g., that includes R (polyadenylation tail signal) and/or U5 which includes a primer activation signal), a primer binding site, a Psi packaging signal, a RRE element for nuclear export, a promoter directly upstream of the transgene to control transgene expression, a transgene (or other exogenous agent element), a polypurine tract, and a 3' LTR (e.g., that includes a mutated U3, a R, and U5). In some embodiments, the retroviral nucleic acid further comprises one or more of a cPPT, a WPRE, and/or an insulator element.

A retrovirus typically replicates by reverse transcription of its genomic RNA into a linear double-stranded DNA copy and subsequently covalently integrates its genomic DNA into a host genome. Illustrative retroviruses suitable for use in particular embodiments, include, but are not limited to: Moloney murine leukemia virus (M-MuLV), Moloney murine sarcoma virus (MoMSV), Harvey murine sarcoma virus (HaMuSV), murine mammary tumor virus (MuMTV), gibbon ape leukemia virus (GaLV), feline leukemia virus (FLV), spumavirus, Friend murine leukemia virus, Murine Stem Cell Virus (MSCV), Rous Sarcoma Virus (RSV), and other lentiviruses.

In some embodiments the retrovirus is a Gammaretrovirus. In some embodiments the retrovirus is an Epsilonretrovirus. In some embodiments the retrovirus is an Alpharetrovirus. In some embodiments the retrovirus is a Betaretrovirus. In some embodiments the retrovirus is a Deltaretrovirus. In some embodiments the retrovirus is a Lentivirus. In some embodiments the retrovirus is a Spumaretrovirus. In some embodiments the retrovirus is an endogenous retrovirus.

Illustrative lentiviruses include, but are not limited to: HIV (human immunodeficiency virus; including HIV type 1, and HIV type 2); visna-maedi virus (VMV) virus; the caprine arthritis-encephalitis virus (CAEV); equine infectious anemia virus (EIAV); feline immunodeficiency virus (FIV); bovine immune deficiency virus (BIV); and simian immunodeficiency virus (SIV). In some embodiments, HIV based vector backbones (i.e., HIV cis-acting sequence elements) are used.

In some embodiments, a vector herein is a nucleic acid molecule capable transferring or transporting another nucleic acid molecule. The transferred nucleic acid is generally linked to, e.g., inserted into, the vector nucleic acid molecule. A vector may include sequences that direct autonomous replication in a cell, or may include sequences sufficient to allow integration into host cell DNA. Useful vectors include, for example, plasmids (e.g., DNA plasmids or RNA plasmids), transposons, cosmids, bacterial artificial chromosomes, and viral vectors. Useful viral vectors include, e.g., replication defective retroviruses and lentiviruses.

In some embodiments, a viral vector comprises a nucleic acid molecule (e.g., a transfer plasmid) that includes virus-derived nucleic acid elements that typically facilitate transfer of the nucleic acid molecule or integration into the genome of a cell or to a viral particle that mediates nucleic acid transfer. Viral particles will typically include various viral components and sometimes also host cell components in addition to nucleic acid(s). In some embodiments, a viral vector comprises e.g., a virus or viral particle capable of transferring a nucleic acid into a cell, or the transferred nucleic acid (e.g., as naked DNA). In some embodiments, a viral vectors and transfer plasmids comprise structural and/or functional genetic elements that are primarily derived from a virus. A retroviral vector can comprise a viral vector or plasmid containing structural and functional genetic elements, or portions thereof, that are primarily derived from a retrovirus. A lentiviral vector can comprise a viral

vector or plasmid containing structural and functional genetic elements, or portions thereof, including LTRs that are primarily derived from a lentivirus.

In embodiments, a lentiviral vector (e.g., lentiviral expression vector) may comprise a lentiviral transfer plasmid (e.g., as naked DNA) or an infectious lentiviral particle.

With respect to elements such as cloning sites, promoters, regulatory elements, heterologous nucleic acids, etc., it is to be understood that the sequences of these elements can be present in RNA form in lentiviral particles and can be present in DNA form in DNA plasmids.

In some embodiments, in the vectors described herein at least part of one or more protein coding regions that contribute to or are essential for replication may be absent compared to the corresponding wild-type virus. In some embodiments, the viral vector is replication-defective. In some embodiments, the vector is capable of transducing a target non-dividing host cell and/or integrating its genome into a host genome.

In some embodiments, different cells differ in their usage of particular codons. In some embodiments, this codon bias corresponds to a bias in the relative abundance of particular tRNAs in the cell type. In some embodiments, by altering the codons in the sequence so that they are tailored to match with the relative abundance of corresponding tRNAs, it is possible to increase expression. In some embodiments, it is possible to decrease expression by deliberately choosing codons for which the corresponding tRNAs are known to be rare in the particular cell type. In some embodiments, an additional degree of translational control is available. An additional description of codon optimization is found, e.g., in WO 99/41397, which is herein incorporated by reference in its entirety.

Conventional techniques for generating retrovirus vectors (and, in particular, lentivirus vectors) with or without the use of packaging/helper vectors are known to those skilled in the art and may be used to generate targeted lipid particles according to the present disclosure. (See, e.g., Derse and Newbold 1993 *Virology* 194:530-6; Maury et al. 1994 *Virology* 200:632-42; Wanisch et al. 2009. *Mol Ther.* 1798:1316-1332; Martarano et al. 1994 *J. Virol.* 68:3102-11; Naldini et al., (1996a, 1996b, and 1998); Zufferey et al., 1999, *J. Virol.*, 73:2886; Huang et al., *Mol. Cell. Biol.*, 5:3864; Liu et al., 1995, *Genes Dev.*, 9:1766; Cullen et al., 1991. *J. Virol.* 65: 1053; and Cullen et al., 1991. *Cell* 58: 423; Dull et al., 1998, U.S. Pat. Nos. 6,013,516; and

5,994,136; PCT patent applications WO 99/15683, WO 98/17815, WO 99/32646, and WO 01/79518). Conventional techniques relating to packaging vectors and producer cells known in the art may also be used according to the present disclosure. (See, e.g., Yao et al, 1998; Jones et al, 2005.)

Provided herein are targeted lipid particles that comprise a naturally derived membrane. In some embodiments, the naturally derived membrane comprises membrane vesicles prepared from cells or tissues. In some embodiments, the targeted lipid particle comprises a vesicle that is obtainable from a cell. In some embodiments, the targeted lipid particle comprises a microvesicle, an exosome, a membrane enclosed body, an apoptotic body (from apoptotic cells), a particle (which may be derived from e.g. platelets), an ectosome (derivable from, e.g., neutrophils and monocytes in serum), a prostasome (obtainable from prostate cancer cells), or a cardiosome (derivable from cardiac cells).

In some embodiments, the source cell is an endothelial cell, a fibroblast, a blood cell (e.g., a macrophage, a neutrophil, a granulocyte, a leukocyte), a stem cell (e.g., a mesenchymal stem cell, an umbilical cord stem cell, bone marrow stem cell, a hematopoietic stem cell, an induced pluripotent stem cell e.g., an induced pluripotent stem cell derived from a subject's cells), an embryonic stem cell (e.g., a stem cell from embryonic yolk sac, placenta, umbilical cord, fetal skin, adolescent skin, blood, bone marrow, adipose tissue, erythropoietic tissue, hematopoietic tissue), a myoblast, a parenchymal cell (e.g., hepatocyte), an alveolar cell, a neuron (e.g., a retinal neuronal cell), a precursor cell (e.g., a retinal precursor cell, a myeloblast, myeloid precursor cells, a thymocyte, a megakaryocyte, a megakaryoblast, a promegakaryoblast, a melanoblast, a lymphoblast, a bone marrow precursor cell, a normoblast, or an angioblast), a progenitor cell (e.g., a cardiac progenitor cell, a satellite cell, a radial glial cell, a bone marrow stromal cell, a pancreatic progenitor cell, an endothelial progenitor cell, a blast cell), or an immortalized cell (e.g., HeEa, HEK293, HFF-I, MRC-5, WI-38, IMR 90, IMR 91, PER.C6, HT-1080, or BJ cell). In some embodiments, the source cell is other than a 293 cell, HEK cell, human endothelial cell, or a human epithelial cell, monocyte, macrophage, dendritic cell, or stem cell.

In some embodiments, the targeted lipid particle has a density of <1, 1-1.1, 1.05-1.15, 1.1-1.2, 1.15-1.25, 1.2-1.3, 1.25-1.35, or >1.35 g/ml. In some embodiments, the

targeted lipid particle composition comprises less than 0.01%, 0.05%, 0.1%, 0.5%, 1%, 1.5%, 2%, 2.5%, 3%, 4%, 5%, or 10% source cells by protein mass, or less than 0.01%, 0.05%, 0.1%, 0.5%, 1%, 1.5%, 2%, 2.5%, 3%, 4%, 5%, or 10% of cells having a functional nucleus.

In embodiments, the targeted lipid particle has a size, or the population of targeted lipid particles have an average size, that is less than about 0.01%, 0.05%, 0.1%, 0.5%, 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, of that of the source cell.

In some embodiments the targeted lipid particle comprises an extracellular vesicle, e.g., a cell-derived vesicle comprising a membrane that encloses an internal space and has a smaller diameter than the cell from which it is derived. In embodiments the extracellular vesicle has a diameter from 20 nm to 1000 nm. In embodiments the targeted lipid particle comprises an apoptotic body, a fragment of a cell, a vesicle derived from a cell by direct or indirect manipulation, a vesiculated organelle, and a vesicle produced by a living cell (e.g., by direct plasma membrane budding or fusion of the late endosome with the plasma membrane). In embodiments the extracellular vesicle is derived from a living or dead organism, explanted tissues or organs, or cultured cells.

In embodiments, the targeted lipid particle comprises a nanovesicle, e.g., a cell-derived small (e.g., between 20-250 nm in diameter, or 30-150 nm in diameter) vesicle comprising a membrane that encloses an internal space, and which is generated from said cell by direct or indirect manipulation. The production of nanovesicles can, in some instances, result in the destruction of the source cell. The nanovesicle may comprise a lipid or fatty acid and polypeptide.

In embodiments, the targeted lipid particle comprises an exosome. In embodiments, the exosome is a cell-derived small (e.g., between 20-300 nm in diameter, or 40-200 nm in diameter) vesicle comprising a membrane that encloses an internal space, and which is generated from said cell by direct plasma membrane budding or by fusion of the late endosome with the plasma membrane. In embodiments, production of exosomes does not result in the destruction of the source cell. In embodiments, the exosome comprises lipid or fatty acid and polypeptide.

In some embodiments, the targeted lipid particle is derived from a source cell with a genetic modification which results in increased expression of an immunomodulatory

agent. In some embodiments, the immunosuppressive agent is on an exterior surface of the cell. In some embodiments, the immunosuppressive agent is incorporated into the exterior surface of the targeted lipid particle. In some embodiments, the targeted lipid particle comprises an immunomodulatory agent attached to the surface of the solid particle by a covalent or non-covalent bond.

a. Generation of cell-derived particles

In some embodiments, targeted lipid particles are generated by inducing budding of an exosome, microvesicle, membrane vesicle, extracellular membrane vesicle, plasma membrane vesicle, giant plasma membrane vesicle, apoptotic body, mitoparticle, pyrenocyte, lysosome, or other membrane enclosed vesicle.

In some embodiments, targeted lipid particles are generated by inducing cell enucleation. Enucleation may be performed using assays such as genetic, chemical (e.g., using Actinomycin D, see Bayona-Bafaluyet al., "A chemical enucleation method for the transfer of mitochondrial DNA to ρ^0 cells" *Nucleic Acids Res.* 2003 Aug 15; 31(16): e98), or mechanical methods (e.g., squeezing or aspiration, see Lee et al., "A comparative study on the efficiency of two enucleation methods in pig somatic cell nuclear transfer: effects of the squeezing and the aspiration methods." *Anim Biotechnol.* 2008;19(2):71-9), or combinations thereof.

In some embodiments, the targeted lipid particles are generated by inducing cell fragmentation. In some embodiments, cell fragmentation can be performed using the following methods, including, but not limited to: chemical methods, mechanical methods (e.g., centrifugation (e.g., ultracentrifugation, or density centrifugation), freeze-thaw, or sonication), or combinations thereof.

In some embodiments, the targeted lipid particle is a microvesicle. In some embodiments the microvesicle has a diameter of about 100 nm to about 2000 nm. In some embodiments, a targeted lipid particle comprises a cell ghost. In some embodiments, a vesicle is a plasma membrane vesicle, e.g., a giant plasma membrane vesicle.

In some embodiments, a characteristic of a targeted lipid particle is described by comparison to a reference cell. In embodiments, the reference cell is the source cell. In embodiments, the reference cell is a HeLa, HEK293, HFF-1, MRC-5, WI-38, IMR 90, IMR 91, PER.C6, HT-1080, or BJ cell. In some embodiments, for example when

the source cell used to make the targeted lipid particle is not available for testing after the targeted lipid particle is made, a characteristic of a population of targeted lipid particle is described by comparison to a population of reference cells, e.g., a population of source cells, or a population of HeLa, HEK293, HFF-1, MRC-5, WI-38, IMR 90, IMR 91, PER.C6, HT-1080, or BJ cells.

Pharmaceutical Compositions

The present disclosure also provides, in some aspects, a pharmaceutical composition comprising the targeted lipid particle (e.g., targeted viral vectors) composition described herein and a pharmaceutically acceptable carrier. The pharmaceutical compositions can include any of the described targeted viral vectors. In some embodiments, the targeted viral vector meets a pharmaceutical or good manufacturing practices (GMP) standard. In some embodiments, the targeted viral vector is made according to good manufacturing practices (GMP). In some embodiments, the targeted viral vector has a pathogen level below a predetermined reference value, e.g., is substantially free of pathogens. In some embodiments, the targeted viral vector has a contaminant level below a predetermined reference value, e.g., is substantially free of contaminants. In some embodiments, the targeted viral vector has low immunogenicity.

In some embodiments, provided herein are the use of pharmaceutical compositions to practice the methods of the disclosure. Such a pharmaceutical composition may consist of at least one targeted lipid particle of the disclosure in a form suitable for administration to a subject, or the pharmaceutical composition may comprise at least one targeted lipid particle of the disclosure and one or more pharmaceutically acceptable carriers, one or more additional ingredients, or some combination of these.

In some embodiments, the relative amounts of the targeted lipid particle, the pharmaceutically acceptable carrier, and any additional ingredients in a pharmaceutical composition of the disclosure will vary, depending upon the identity, size, and condition of the subject treated and further depending upon the route by which the composition is to be administered. In some embodiments, the composition may comprise between 0.1% and 100% (w/w) of the targeted lipid particles of the disclosure.

In some embodiments, pharmaceutical compositions that are useful in the methods of the disclosure may be suitably developed for intravenous, intratumoral, oral, rectal, vaginal, parenteral, topical, pulmonary, intranasal, buccal, ophthalmic, or another route of administration. In some embodiments, a composition useful within the methods of the disclosure may be directly administered to the skin, vagina or any other tissue of a mammal. In some embodiments, formulations include liposomal preparations, resealed erythrocytes containing the targeted lipid particles of the disclosure, and immunologically based formulations. In some embodiments, the route(s) of administration will be readily apparent to the skilled artisan and will depend upon any number of factors including the type and severity of the disease being treated, the type and age of the veterinary or human subject being treated, and the like.

In some embodiments, formulations of the pharmaceutical compositions described herein may be prepared by any method known or hereafter developed in the art of pharmacology. In some embodiments, preparatory methods include the step of bringing the targeted lipid particles of the disclosure into association with a carrier or one or more other accessory ingredients, and then, if necessary or desirable, shaping or packaging the product into a desired single- or multi-dose unit.

In some embodiments, a "unit dose" is a discrete amount of the pharmaceutical composition comprising a predetermined amount of the targeted lipid particles of the disclosure. In some embodiments, the amount is generally equal to the dosage that would be administered to a subject or a convenient fraction of such a dosage such as, for example, one-half or one-third of such a dosage. In some embodiments, the unit dosage form may be for a single daily dose or one of multiple daily doses (e.g., about 1 to 4 or more times per day). In some embodiments, when multiple daily doses are used, the unit dosage form may be the same or different for each dose. In some embodiments, although the descriptions of pharmaceutical compositions provided herein are principally directed to pharmaceutical compositions that are suitable for ethical administration to humans, it will be understood by the skilled artisan that such compositions are generally suitable for administration to animals of all sorts. In some embodiments, modification of pharmaceutical compositions suitable for administration to humans in order to render the compositions suitable for administration to various animals is well understood, and the ordinarily skilled

veterinary pharmacologist may design and perform such modification with merely ordinary, if any, experimentation. In some embodiments, subjects to which administration of the pharmaceutical compositions of the disclosure is contemplated include humans and other primates, mammals including commercially relevant mammals such as cattle, pigs, horses, sheep, cats, and dogs.

In some of any embodiments, the compositions of the disclosure are formulated using one or more pharmaceutically acceptable excipients or carriers. In one embodiment, the pharmaceutical compositions of the disclosure comprise a therapeutically effective amount of a targeted lipid particle of the disclosure and a pharmaceutically acceptable carrier. In some embodiments, pharmaceutically acceptable carriers that are useful, include, but are not limited to, glycerol, water, saline, ethanol, and other pharmaceutically acceptable salt solutions such as phosphates and salts of organic acids. Examples of these and other pharmaceutically acceptable carriers are described in Remington's Pharmaceutical Sciences (1991, Mack Publication Co., New Jersey).

In some embodiments, the carrier may be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), suitable mixtures thereof, and vegetable oils. In some embodiments, the proper fluidity may be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. In some embodiments, prevention of the action of microorganisms may be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In some embodiments, it is preferable to include isotonic agents, for example, sugars, sodium chloride, or polyalcohols such as mannitol and sorbitol, in the composition. In some embodiments, prolonged absorption of the injectable compositions may be brought about by including in the composition an agent that delays absorption, for example, aluminum monostearate or gelatin. In one embodiment, the pharmaceutically acceptable carrier is not DMSO alone.

In some embodiments, formulations may be employed in admixtures with conventional excipients, i.e., pharmaceutically acceptable organic or inorganic carrier substances suitable for oral, vaginal, parenteral, nasal, intravenous, subcutaneous,

enteral, or any other suitable mode of administration, known to the art. In some embodiments, the pharmaceutical preparations may be sterilized and, if desired, mixed with auxiliary agents, e.g., lubricants, preservatives, stabilizers, wetting agents, emulsifiers, salts for influencing osmotic pressure buffers, coloring, flavoring, and/or aromatic substances and the like. In some embodiments, pharmaceutical preparations may also be combined with other active agents, e.g., other analgesic agents.

In some embodiments, "additional ingredients" include, but are not limited to, one or more of the following: excipients; surface active agents; dispersing agents; inert diluents; granulating and disintegrating agents; binding agents; lubricating agents; sweetening agents; flavoring agents; coloring agents; preservatives; physiologically degradable compositions such as gelatin; aqueous vehicles and solvents; oily vehicles and solvents; suspending agents; dispersing or wetting agents; emulsifying agents, demulcents; buffers; salts; thickening agents; fillers; emulsifying agents; antioxidants; antibiotics; antifungal agents; stabilizing agents; and pharmaceutically acceptable polymeric or hydrophobic materials. In some embodiments, "additional ingredients" that may be included in the pharmaceutical compositions of the disclosure are known in the art and described, for example in Genaro, ed. (1985, Remington's Pharmaceutical Sciences, Mack Publishing Co., Easton, Pa.), which is incorporated herein by reference.

In some embodiments, the composition of the disclosure may comprise a preservative from about 0.005% to 2.0% by total weight of the composition. In some embodiments, the preservative is used to prevent spoilage in the case of exposure to contaminants in the environment. In some embodiments, examples of preservatives useful in accordance with the disclosure included but are not limited to those selected from the group consisting of benzyl alcohol, sorbic acid, parabens, imidurea and combinations thereof. In some embodiments, a particularly preferred preservative is a combination of about 0.5% to 2.0% benzyl alcohol and 0.05% to 0.5% sorbic acid.

In some embodiments, liquid suspensions may be prepared using conventional methods to achieve suspension of the targeted lipid particles of the disclosure in an aqueous or oily vehicle. In some embodiments, aqueous vehicles include, for example, water, and isotonic saline. In some embodiments, oily vehicles include, for

example, almond oil, oily esters, ethyl alcohol, vegetable oils such as arachis, olive, sesame, or coconut oil, fractionated vegetable oils, and mineral oils such as liquid paraffin. In some embodiments, liquid suspensions may further comprise one or more additional ingredients including, but not limited to, suspending agents, dispersing or wetting agents, emulsifying agents, demulcents, preservatives, buffers, salts, flavorings, coloring agents, and sweetening agents. In some embodiments, oily suspensions may further comprise a thickening agent. In some embodiments, suspending agents include, but are not limited to, sorbitol syrup, hydrogenated edible fats, sodium alginate, polyvinylpyrrolidone, gum tragacanth, gum acacia, and cellulose derivatives such as sodium carboxymethylcellulose, methylcellulose, hydroxypropylmethylcellulose. In some embodiments, dispersing or wetting agents include, but are not limited to, naturally-occurring phosphatides such as lecithin, condensation products of an alkylene oxide with a fatty acid, with a long chain aliphatic alcohol, with a partial ester derived from a fatty acid and a hexitol, or with a partial ester derived from a fatty acid and a hexitol anhydride (e.g., polyoxyethylene stearate, heptadecaethyleneoxycetanol, polyoxyethylene sorbitol monooleate, and polyoxyethylene sorbitan monooleate, respectively). Known emulsifying agents include, but are not limited to, lecithin, and acacia. Known preservatives include, but are not limited to, methyl, ethyl, or n-propyl-para-hydroxybenzoates, ascorbic acid, and sorbic acid. Known sweetening agents include, for example, glycerol, propylene glycol, sorbitol, sucrose, and saccharin. Known thickening agents for oily suspensions include, for example, beeswax, hard paraffin, and cetyl alcohol. In some embodiments, liquid solutions of the targeted lipid particles of the disclosure in aqueous or oily solvents may be prepared in substantially the same manner as liquid suspensions, the primary difference being that the targeted lipid particles of the disclosure is dissolved, rather than suspended in the solvent. As used herein, an "oily" liquid is one which comprises a carbon-containing liquid molecule and which exhibits a less polar character than water. In some embodiments, liquid solutions of the pharmaceutical composition of the disclosure may comprise each of the components described with regard to liquid suspensions, it being understood that suspending agents will not necessarily aid dissolution of the targeted lipid particles of the disclosure in the solvent. In some embodiments, aqueous solvents include, for example, water, and isotonic saline. In some

embodiments, oily solvents include, for example, almond oil, oily esters, ethyl alcohol, vegetable oils such as arachis, olive, sesame, or coconut oil, fractionated vegetable oils, and mineral oils such as liquid paraffin.

In some embodiments, powdered and granular formulations of a pharmaceutical preparation of the disclosure may be prepared using known methods. In some embodiments, formulations may be administered directly to a subject, used, for example, to form tablets, to fill capsules, or to prepare an aqueous or oily suspension or solution by addition of an aqueous or oily vehicle thereto. In some of any embodiments, formulations may further comprise one or more of dispersing or wetting agent, a suspending agent, and a preservative. Additional excipients, such as fillers and sweetening, flavoring, or coloring agents, may also be included in these formulations.

In some embodiments, a pharmaceutical composition of the disclosure may also be prepared, packaged, or sold in the form of oil-in-water emulsion or a water-in-oil emulsion. In some embodiments, the oily phase may be a vegetable oil such as olive or arachis oil, a mineral oil such as liquid paraffin, or a combination of these. In some embodiments, compositions further comprise one or more emulsifying agents such as naturally occurring gums such as gum acacia or gum tragacanth, naturally-occurring phosphatides such as soybean or lecithin phosphatide, esters or partial esters derived from combinations of fatty acids and hexitol anhydrides such as sorbitan monooleate, and condensation products of such partial esters with ethylene oxide such as polyoxyethylene sorbitan monooleate. In some embodiments, emulsions may also contain additional ingredients including, for example, sweetening or flavoring agents.

Methods of Treatment

Provided herein are methods of administering a lentiviral vector comprising a CD8 binding agent to a subject. In some embodiments the method comprises a) obtaining whole blood from the subject; b) collecting the fraction of blood containing leukocyte components including CD8+ T cells; c) contacting the leukocyte components including CD8+T cells with a composition comprising the lentiviral vector to create a transfection mixture; and d) reinfusing the contacted leukocyte components including CD8+T cells and/or the transfection mixture to the subject, thereby administering the

lipid particle and/or payload gene to the subject. In some embodiments, the T cells (e.g. CD8+T cells) are not activated during the method. In some embodiments, step (c) of the method is carried out for no more than 24 hours, e.g., no more than 20, 16, 12, 8, 6, 5, 4, 3, 2, or 1 hour.

In some embodiments, the method according to the present disclosure is capable of delivering a lentiviral particle to an ex vivo system. The method includes the use of a combination of various apheresis machine hardware components, a software control module, and a sensor module to measure citrate or other solute levels in-line to ensure the maximum accuracy and safety of treatment prescriptions, and the use of replacement fluids designed to fully exploit the design of the system according to the present methods. It is understood that components described for one system according to the present invention can be implemented within other systems according to the present invention as well.

In some embodiments, the method for administration of the lentiviral vector to the subject comprises the use of a blood processing set for obtaining whole blood from the subject, a separation chamber for collecting the fraction of blood containing leukocyte components including CD8+ T cells, a contacting container for contacting the CD8+T cells with the composition comprising the lentiviral vector, and a further fluid circuit for reinfusion of CD8+ T cells to the patient. In some embodiments, the method further comprises any of i) a washing component for concentrating T cells, and ii) a sensor and/or module for monitoring cell density and/or concentration. In some embodiments, the methods allow processing of blood directly from the patient, transduction with the lentiviral vector, and reinfusion directly to the patient without any steps of selection for T cells or for CD8+T cells. Further, in some embodiments the methods are carried out without cryopreserving or freezing any cells before or between any one or more of the steps, such that there is no step of formulating cells with a cryoprotectant, e.g. DMSO. In some embodiments, the provided methods do not include a lymphodepletion regimen. In some embodiments, the method including steps (a)-(d) are carried out for a time of no more than 24 hours, such as between 2 hours and 12 hours, for example 3 hours to 6 hours.

In some embodiments, the method is performed in-line (or in situ). In some embodiments, the method is performed in a closed fluid circuit, or a functionally closed fluid circuit. In some embodiments, each of steps (a)-(d) are performed in-line

in a closed fluid circuit in which all parts of the system are operably connected, such as via at least one tubing line. In some embodiments, the system is sterile. In some embodiments, the closed fluid circuit is sterile.

Also provided herein are systems for administration of a lentiviral vector comprising a CD8 binding agent to a subject. An exemplary system for administration is shown in **Fig. 14**.

In some embodiments, the targeted lipid particles (e.g. targeted viral vectors) provided herein, or pharmaceutical compositions thereof as described herein can be administered to a subject, e.g. a mammal, e.g. a human. In such embodiments, the subject may be at risk of, may have a symptom of, or may be diagnosed with or identified as having, a particular disease or condition. In one embodiment, the subject has cancer. In one embodiment, the subject has an infectious disease. In some embodiments, the targeted viral vector contains nucleic acid sequences encoding an exogenous agent for treating the disease or condition in the subject. For example, the exogenous agent is one that targets or is specific for a protein of a neoplastic cells and the targeted lipid particle is administered to a subject for treating a tumor or cancer in the subject. In another example, the exogenous agent is an inflammatory mediator or immune molecule, such as a cytokine, and targeted lipid particle is administered to a subject for treating any condition in which it is desired to modulate (e.g., increase) the immune response, such as a cancer or infectious disease. In some embodiments, the targeted viral vector is administered in an effective amount or dose to effect treatment of the disease, condition or disorder. Provided herein are uses of any of the provided targeted viral vector in such methods and treatments, and in the preparation of a medicament in order to carry out such therapeutic methods. In some embodiments, the methods are carried out by administering the targeted viral vector or compositions comprising the same, to the subject having, having had, or suspected of having the disease or condition or disorder. In some embodiments, the methods thereby treat the disease or condition or disorder in the subject. Also provided herein are uses of any of the compositions, such as pharmaceutical compositions provided herein, for the treatment of a disease, condition or disorder associated with a particular gene or protein targeted by or provided by the exogenous agent.

In some embodiments, the provided methods or uses involve administration of a pharmaceutical composition comprising oral, inhaled, transdermal or parenteral (including intravenous, intratumoral, intraperitoneal, intramuscular, intracavity, and subcutaneous) administration. In some embodiments, the targeted viral vector may be administered alone or formulated as a pharmaceutical composition. In some embodiments, the targeted viral vector or compositions described herein can be administered to a subject, e.g., a mammal, e.g., a human. In some of any embodiments, the subject may be at risk of, may have a symptom of, or may be diagnosed with or identified as having, a particular disease or condition (e.g., a disease or condition described herein). In some embodiments, the disease is a disease or disorder. In some embodiments, the disease is a B cell malignancy. In some embodiments, the targeted viral vectors may be administered in the form of a unit-dose composition, such as a unit dose oral, parenteral, transdermal, or inhaled composition. In some embodiments, the compositions are prepared by admixture and are adapted for oral, inhaled, transdermal, or parenteral administration, and as such may be in the form of tablets, capsules, oral liquid preparations, powders, granules, lozenges, reconstitutable powders, injectable, and infusable solutions or suspensions, or suppositories or aerosols.

In some embodiments, the regimen of administration may affect what constitutes an effective amount. In some embodiments, the therapeutic formulations may be administered to the subject either prior to or after a diagnosis of disease. In some embodiments, several divided dosages, as well as staggered dosages may be administered daily or sequentially, or the dose may be continuously infused, or may be a bolus injection. In some embodiments, the dosages of the therapeutic formulations may be proportionally increased or decreased as indicated by the exigencies of the therapeutic or prophylactic situation.

In some embodiments, the administration of the compositions of the present disclosure to a subject, preferably a mammal, more preferably a human, may be carried out using known procedures, at dosages and for periods of time effective to prevent or treat disease. In some embodiments, an effective amount of the targeted lipid particle of the disclosure necessary to achieve a therapeutic effect may vary according to factors such as the activity of the particular lipid particle employed; the time of administration; the rate of excretion; the duration of the treatment; other

drugs, compounds or materials used in combination with the targeted lipid particle of the disclosure; the state of the disease or disorder, age, sex, weight, condition, general health and prior medical history of the subject being treated, and like factors well-known in the medical arts. In some embodiments, the dosage regimens may be adjusted to provide the optimum therapeutic response. In some embodiments, several divided doses may be administered daily or the dose may be proportionally reduced as indicated by the exigencies of the therapeutic situation. One of ordinary skill in the art would be able to study the relevant factors and make the determination regarding the effective amount of the therapeutic targeted lipid particle of the disclosure without undue experimentation.

In some embodiments, dosage levels of the targeted lipid particles in the pharmaceutical compositions of this disclosure may be varied so as to obtain an amount that is effective to achieve the desired therapeutic response for a particular subject, composition, and mode of administration, without being toxic to the subject. A medical doctor, e.g., physician or veterinarian, having ordinary skill in the art may readily determine and prescribe the effective amount of the pharmaceutical composition required. In some embodiments, the physician or veterinarian could start doses of the targeted lipid particles of the disclosure employed in the pharmaceutical composition at levels lower than that required in order to achieve the desired therapeutic effect and gradually increase the dosage until the desired effect is achieved.

In some embodiments, the term "container" includes any receptacle for holding the pharmaceutical composition. In some embodiments, the container is the packaging that contains the pharmaceutical composition. In other embodiments, the container is not the packaging that contains the pharmaceutical composition, i.e., the container is a receptacle, such as a box or vial that contains the packaged pharmaceutical composition or unpackaged pharmaceutical composition and the instructions for use of the pharmaceutical composition. It should be understood that the instructions for use of the pharmaceutical composition may be contained on the packaging containing the pharmaceutical composition, and as such the instructions form an increased functional relationship to the packaged product. In some embodiments, instructions may contain information pertaining to the pharmaceutical composition's

ability to perform its intended function, e.g., treating or preventing a disease in a subject, or delivering an imaging or diagnostic agent to a subject.

In some embodiments, routes of administration of any of the compositions disclosed herein include oral, nasal, rectal, parenteral, sublingual, transdermal, transmucosal (e.g., sublingual, lingual, (trans)buccal, (trans)urethral, vaginal (e.g., trans- and perivaginally), (intra)nasal, and (trans)rectal), intravesical, intrapulmonary, intraduodenal, intragastrical, intrathecal, subcutaneous, intramuscular, intradermal, intra-arterial, intravenous, intrabronchial, inhalation, and topical administration.

In some of any embodiments, suitable compositions and dosage forms include, for example, tablets, capsules, caplets, pills, gel caps, troches, dispersions, suspensions, solutions, syrups, granules, beads, transdermal patches, gels, powders, pellets, magmas, lozenges, creams, pastes, plasters, lotions, discs, suppositories, liquid sprays for nasal or oral administration, dry powder or aerosolized formulations for inhalation, compositions and formulations for intravesical administration, and the like.

In some embodiments, the targeted lipid particle composition comprising an exogenous agent or cargo, may be used to deliver such exogenous agent or cargo to a cell tissue or subject. In some embodiments, delivery of a cargo by administration of a targeted lipid particle composition described herein may modify cellular protein expression levels. In certain embodiments, the administered composition directs upregulation (via expression in the cell, delivery in the cell, or induction within the cell) of one or more cargo (e.g., a polypeptide or mRNA) that provide a functional activity which is substantially absent or reduced in the cell in which the polypeptide is delivered. In some embodiments, the missing functional activity may be enzymatic, structural, or regulatory in nature. In some embodiments, the administered composition directs up-regulation of one or more polypeptides that increases (e.g., synergistically) a functional activity which is present but substantially deficient in the cell in which the polypeptide is upregulated. In some of any embodiments, the administered composition directs downregulation of (via expression in the cell, delivery in the cell, or induction within the cell) of one or more cargo (e.g., a polypeptide, siRNA, or miRNA) that repress a functional activity which is present or upregulated in the cell in which the polypeptide, siRNA, or miRNA is delivered. In some embodiments, the upregulated functional activity may be

enzymatic, structural, or regulatory in nature. In some embodiments, the administered composition directs down-regulation of one or more polypeptides that decreases (e.g., synergistically) a functional activity which is present or upregulated in the cell in which the polypeptide is downregulated. In some embodiments, the administered composition directs upregulation of certain functional activities and downregulation of other functional activities.

In some of any embodiments, the targeted lipid particle composition (e.g., one comprising mitochondria or DNA) mediates an effect on a target cell, and the effect lasts for at least 1, 2, 3, 4, 5, 6, or 7 days, 2, 3, or 4 weeks, or 1, 2, 3, 6, or 12 months. In some embodiments (e.g., wherein the targeted viral vector composition comprises an exogenous protein), the effect lasts for less than 1, 2, 3, 4, 5, 6, or 7 days, 2, 3, or 4 weeks, or 1, 2, 3, 6, or 12 months.

In some of any embodiments, the targeted lipid particle composition described herein is delivered ex-vivo to a cell or tissue, e.g., a human cell or tissue. In embodiments, the composition improves function of a cell or tissue ex-vivo, e.g., improves cell viability, respiration, or other function (e.g., another function described herein).

In some embodiments, the composition is delivered to an ex vivo tissue that is in an injured state (e.g., from trauma, disease, hypoxia, ischemia or other damage).

In some embodiments, the composition is delivered to an ex-vivo transplant (e.g., a tissue explant or tissue for transplantation, e.g., a human vein, a musculoskeletal graft such as bone or tendon, cornea, skin, heart valves, nerves; or an isolated or cultured organ, e.g., an organ to be transplanted into a human, e.g., a human heart, liver, lung, kidney, pancreas, intestine, thymus, eye). In some embodiments, the composition is delivered to the tissue or organ before, during and/or after transplantation.

In some embodiments, the composition is delivered, administered, or contacted with a cell, e.g., a cell preparation. In some embodiments, the cell preparation may be a cell therapy preparation (a cell preparation intended for administration to a human subject). In embodiments, the cell preparation comprises cells expressing a T-cell receptor (TCR) or chimeric antigen receptor (CAR), e.g., expressing a recombinant CAR. The cells expressing the CAR may be, e.g., T cells, Natural Killer (NK) cells, cytotoxic T lymphocytes (CTL), regulatory T cells. In embodiments, the cell preparation is a neural stem cell preparation. In embodiments, the cell preparation is

a mesenchymal stem cell (MSC) preparation. In embodiments, the cell preparation is a hematopoietic stem cell (HSC) preparation. In embodiments, the cell preparation is an islet cell preparation.

In some embodiments, the viral vector comprising an anti-CD8 sdAb or scFv composition described herein is used to deliver a CAR or TCR. In some embodiments, the viral vector transduces a cell expressing CD8 (e.g., a CD8⁺ T cell) and expresses and amplifies the CAR or TCR. The amplified CAR or TCR T cells then mediate targeted cell killing. Thus, the disclosure includes the use of viral vector comprising an anti-CD8 scFv fusogen construct to elicit an immune response specific to the antigen binding moiety of the CAR or TCR. In some embodiments, the CAR is used to target a tumor antigen selected from CD19, CD20, CD22, or BCMA. In another embodiment, the CAR is engineered to comprise an intracellular signaling domain of the T cell antigen receptor complex zeta chain (e.g., CD3 zeta). In a preferred embodiment, the intracellular domain is selected from a CD137 (4-1BB) signaling domain, a CD28 signaling domain, and a CD3zeta signaling domain.

Engineered Receptor Payloads

In some embodiments, the targeted lipid particles (e.g. targeted viral vectors) disclosed herein encode an engineered receptor. In some embodiments, the cells for use in or administered in connection with the provided methods contain or are engineered to contain an engineered receptor, e.g., an engineered antigen receptor, such as a chimeric antigen receptor (CAR). Also provided are populations of such cells, compositions containing such cells and/or enriched for such cells, such as in which cells of a certain type such as T cells or CD8⁺ cells are enriched or selected. Among the compositions are pharmaceutical compositions and formulations for administration, such as for adoptive cell therapy. Also provided are therapeutic methods for administering the cells and compositions to subjects, e.g., patients, in accord with the provided methods, and/or with the provided articles of manufacture or compositions.

In some embodiments, gene transfer is accomplished without first stimulating the cells, such as by combining it with a stimulus that induces a response such as proliferation, survival, and/or activation, e.g., as measured by expression of a cytokine or activation marker, followed by introduction of the nucleic acids, e.g., by

transduction, into the stimulated cells, and optionally incubation or expansion in culture to numbers sufficient for clinical applications.

The viral vectors may express recombinant receptors, such as antigen receptors including chimeric antigen receptors (CARs), and other antigen-binding receptors such as transgenic T cell receptors (TCRs). Also among the receptors are other chimeric receptors.

a. Chimeric Antigen Receptors (CARs)

In some embodiments of the provided methods and uses, chimeric receptors, such as a CARs, contain one or more domains that combine an antigen- or ligand-binding domain (e.g. antibody or antibody fragment) that provides specificity for a desired antigen (e.g., tumor antigen) with intracellular signaling domains. In some embodiments, the intracellular signaling domain is a stimulating or an activating intracellular domain portion, such as a T cell stimulating or activating domain, providing a primary activation signal or a primary signal. In some embodiments, the intracellular signaling domain contains or additionally contains a costimulatory signaling domain to facilitate effector functions. In some embodiments, chimeric receptors when genetically engineered into immune cells can modulate T cell activity, and, in some cases, can modulate T cell differentiation or homeostasis, thereby resulting in genetically engineered cells with improved longevity, survival and/or persistence in vivo, such as for use in adoptive cell therapy methods. Exemplary antigen receptors, including CARs, and methods for engineering and introducing such receptors into cells, include those described, for example, in WO200014257, WO2013126726, WO2012/129514, WO2014031687, WO2013/166321, WO2013/071154, WO2013/123061, U.S. patent app. Pub. Nos. US2002131960, US2013287748, US20130149337, U.S. Patent Nos. 6,451,995, 7,446,190, 8,252,592, 8,339,645, 8,398,282, 7,446,179, 6,410,319, 7,070,995, 7,265,209, 7,354,762, 7,446,191, 8,324,353, and 8,479,118, and European patent app. No. EP2537416, and/or those described by Sadelain et al., *Cancer Discov.* 2013 April; 3(4): 388-398; Davila et al. (2013) *PLoS ONE* 8(4): e61338; Turtle et al., *Curr. Opin. Immunol.*, 2012 October; 24(5): 633-39; Wu et al., *Cancer*, 2012 March 18(2): 160-75. In some aspects, the antigen receptors include a CAR as described in U.S. Patent No.: 7,446,190, and those described in WO/2014055668. Examples of

the CARs include CARs as disclosed in any of the aforementioned publications, such as WO2014031687, US 8,339,645, US 7,446,179, US 2013/0149337, US 7,446,190, US 8,389,282, Kochenderfer et al., (2013) Nature Reviews Clinical Oncology, 10, 267-276; Wang et al. (2012) J. Immunother. 35(9): 689-701; and Brentjens et al., Sci Transl Med. 2013 5(177). See also WO2014031687, US 8,339,645, US 7,446,179, US 2013/0149337, US 7,446,190, and US 8,389,282. The recombinant receptors, such as CARs, generally include an extracellular antigen binding domain, such as a portion of an antibody molecule, generally a variable heavy (VH) chain region and/or variable light (VL) chain region of the antibody, e.g., an scFv antibody fragment. In some embodiments, the antigen binding domain of the CAR molecule comprises an antibody, an antibody fragment, an scFv, a Fv, a Fab, a (Fab')₂, a single domain antibody (SdAb), a VH or VL domain, or a camelid VHH domain.

In some embodiments, the antigen targeted by the receptor is a polypeptide. In some embodiments, it is a carbohydrate or other molecule. In some embodiments, the antigen is selectively expressed or overexpressed on cells of the disease or condition, e.g., the tumor or pathogenic cells, as compared to normal or non-targeted cells or tissues. In other embodiments, the antigen is expressed on normal cells and/or is expressed on the engineered cells.

In some embodiments, the antigen targeted by the receptor includes antigens associated with a B cell malignancy, such as any of a number of known B cell marker. In some embodiments, the antigen targeted by the receptor is CD20, CD19, CD22, ROR1, CD45, CD47, CD21, CD5, CD33, Igkappa, Iglambda, CD79a, CD79b or CD30.

In some embodiments, the chimeric antigen receptor includes an extracellular portion containing an antibody or antibody fragment. In some aspects, the chimeric antigen receptor includes an extracellular portion containing the antibody or fragment and an intracellular signaling domain. In some embodiments, the antibody or fragment includes an scFv.

In some embodiments, the antigen targeted by the antigen-binding domain is CD19. In some aspects, the antigen-binding domain of the recombinant receptor, e.g., CAR, and the antigen-binding domain binds, such as specifically binds or specifically recognizes, a CD19, such as a human CD19. In some embodiments, the scFv contains a VH and a VL derived from an antibody or an antibody fragment specific to

CD19. In some embodiments, the antibody or antibody fragment that binds CD19 is a mouse derived antibody such as FMC63 and SJ25C1. In some embodiments, the antibody or antibody fragment is a human antibody, e.g., as described in U.S. Patent Publication No. US 2016/0152723.

In some embodiments, the antigen is CD19. In some embodiments, the scFv contains a VH and a VL derived from an antibody or an antibody fragment specific to CD 19. In some embodiments, the antibody or antibody fragment that binds CD 19 is a mouse derived antibody such as FMC63 and SJ25C1. In some embodiments, the antibody or antibody fragment is a human antibody, e.g., as described in U.S. Patent Publication No. US 2016/0152723.

In some embodiments, the scFv is derived from FMC63. FMC63 generally refers to a mouse monoclonal IgG1 antibody raised against Nalm-1 and -16 cells expressing CD19 of human origin (Fing, N. R., et al. (1987). Leucocyte typing III. 302).

In some embodiments, the antigen targeted by the antigen-binding domain is BCMA. In some aspects, the antigen-binding domain of the recombinant receptor, e.g., CAR, and the antigen-binding domain binds, such as specifically binds or specifically recognizes, a BCMA, such as a human BCMA. In some embodiments, the antigen-binding domain is a fully human VH sdAb disclosed in US2020/0138865 (disclosed herein by reference in its entirety), e.g., FHVH74, FHVH32, FHVH33, or FHVH93.

In some embodiments, the CD19 specific CAR includes an anti-CD19 single-chain antibody fragment (scFv), a transmembrane domain such as one derived from human CD8 α , a 4-1BB (CD137) co-stimulatory signaling domain, and a CD3 ζ signaling domain. In some embodiments, the CD22 specific CAR includes an anti-CD22 scFv, a transmembrane domain such as one derived from human CD8 α , a 4-1BB (CD137) co-stimulatory signaling domain, and a CD3 ζ signaling domain. In some embodiments, the CD19/CD22-bispecific CAR includes an anti-CD19 scFv, an anti-CD22 scFv, a transmembrane domain such as one derived from human CD8 α , a 4-1BB (CD137) co-stimulatory signaling domain, and a CD3 ζ signaling domain.

In some embodiments, the CAR comprises a commercial CAR construct carried by a T cell. Non-limiting examples of commercial CAR-T cell based therapies include brexucabtagene autoleucel (TECARTUS®), axicabtagene ciloleucel (YESCARTA®), idecabtagene vicleucel (ABECMA®), lisocabtagene maraleucel (BREYANZI®), tisagenlecleucel (KYMRIA®), Descartes-08 and Descartes-11 from Cartesian

Therapeutics, CTL110 from Novartis, P-BMCA-101 from Poseida Therapeutics, AUTO4 from Autolus Limited, UCARTCS from Cellectis, PBCAR19B and PBCAR269A from Precision Biosciences, FT819 from Fate Therapeutics, and CYAD-211 from Clyad Oncology.

In some embodiments, a chimeric antigen receptor (CAR) comprises an antigen binding domain. In some embodiments, the CAR is or comprises a first generation CAR comprising an antigen binding domain, a transmembrane domain, and at least one signaling domain (*e.g.*, one, two or three signaling domains). In some embodiments, the CAR comprises a second generation CAR comprising an antigen binding domain, a transmembrane domain, and at least two signaling domains. In some embodiments, the CAR comprises a third generation CAR comprising an antigen binding domain, a transmembrane domain, and at least three signaling domains. In some embodiments, a fourth generation CAR comprising an antigen binding domain, a transmembrane domain, three or four signaling domains, and a domain which upon successful signaling of the CAR induces expression of a cytokine gene. In some embodiments, the antigen binding domain is or comprises an antibody, an antibody fragment, an scFv or a Fab.

1. Antigen binding domain (ABD) targets an antigen characteristic of a neoplastic or cancer cell

In some embodiments, the antigen binding domain (ABD) targets an antigen characteristic of a neoplastic cell. In other words, the antigen binding domain targets an antigen expressed by a neoplastic or cancer cell. In some embodiments, the ABD binds a tumor associated antigen. In some embodiments, the antigen characteristic of a neoplastic cell (*e.g.*, antigen associated with a neoplastic or cancer cell) or a tumor associated antigen is selected from a cell surface receptor, an ion channel-linked receptor, an enzyme-linked receptor, a G protein-coupled receptor, receptor tyrosine kinase, tyrosine kinase associated receptor, receptor-like tyrosine phosphatase, receptor serine/ threonine kinase, receptor guanylyl cyclase, histidine kinase associated receptor, epidermal growth factor receptors (EGFR) (including ErbB1/EGFR, ErbB2/HER2, ErbB3/HER3, and ErbB4/HER4), fibroblast growth factor receptors (FGFR) (including FGF1, FGF2, FGF3, FGF4, FGF5, FGF6, FGF7, FGF18, and FGF21), vascular endothelial growth factor receptors (VEGFR)

(including VEGF-A, VEGF-B, VEGF-C, VEGF-D, and PIGF), RET Receptor and the Eph Receptor Family (including EphA1, EphA2, EphA3, EphA4, EphA5, EphA6, EphA7, EphA8, EphA9, EphA10, EphB1, EphB2, EphB3, EphB4, and EphB6), CXCR1, CXCR2, CXCR3, CXCR4, CXCR6, CCR1, CCR2, CCR3, CCR4, CCR5, CCR6, CCR8, CFTR, CIC-1, CIC-2, CIC-4, CIC-5, CIC-7, CIC-Ka, CIC-Kb, Bestrophins, TMEM16A, GABA receptor, glycin receptor, ABC transporters, NAV1.1, NAV1.2, NAV1.3, NAV1.4, NAV1.5, NAV1.6, NAV1.7, NAV1.8, NAV1.9, sphingosin-1-phosphate receptor (S1P1R), NMDA channel, transmembrane protein, multspan transmembrane protein, T-cell receptor motifs, T-cell alpha chains, T-cell β chains, T-cell γ chains, T-cell δ chains, CCR7, CD3, CD4, CD5, CD7, CD8, CD11b, CD11c, CD16, CD19, CD20, CD21, CD22, CD25, CD28, CD34, CD35, CD40, CD45RA, CD45RO, CD52, CD56, CD62L, CD68, CD80, CD95, CD117, CD127, CD133, CD137 (4-1BB), CD163, F4/80, IL-4Ra, Sca-1, CTLA-4, GITR, GARP, LAP, granzyme B, LFA-1, transferrin receptor, NKp46, perforin, CD4+, Th1, Th2, Th17, Th40, Th22, Th9, Tfh, canonical Treg, FoxP3+, Tr1, Th3, Treg17, T_{REG}; CDCP, NT5E, EpCAM, CEA, gpA33, mucins, TAG-72, carbonic anhydrase IX, PSMA, folate binding protein, gangliosides (e.g., CD2, CD3, GM2), Lewis- γ^2 , VEGF, VEGFR 1/2/3, α V β 3, α 5 β 1, ErbB1/EGFR, ErbB1/HER2, ErbB3, c-MET, IGF1R, EphA3, TRAIL-R1, TRAIL-R2, RANKL, FAP, Tenascin, PDL-1, BAFF, HDAC, ABL, FLT3, KIT, MET, RET, IL-1 β , ALK, RANKL, mTOR, CTLA-4, IL-6, IL-6R, JAK3, BRAF, PTCH, Smoothened, PIGF, ANPEP, TIMP1, PLAUR, PTPRJ, LTBR, ANTXR1, folate receptor alpha (FRa), ERBB2 (Her2/neu), EphA2, IL-13Ra2, epidermal growth factor receptor (EGFR), mesothelin, TSHR, CD19, CD123, CD22, CD30, CD171, CS-1, CLL-1, CD33, EGFRvIII, GD2, GD3, BCMA, MUC16 (CA125), L1CAM, LeY, MSLN, IL13R α 1, L1-CAM, Tn Ag, prostate specific membrane antigen (PSMA), ROR1, FLT3, FAP, TAG72, CD38, CD44v6, CEA, EPCAM, B7H3, KIT, interleukin-11 receptor a (IL-11Ra), PSCA, PRSS21, VEGFR2, LewisY, CD24, platelet-derived growth factor receptor-beta (PDGFR-beta), SSEA-4, CD20, MUC1, NCAM, Prostase, PAP, ELF2M, Ephrin B2, IGF-1 receptor, CAIX, LMP2, gp100, bcr-abl, tyrosinase, Fucosyl GM1, sLe, GM3, TGS5, HMWMAA, o-acetyl-GD2, folate receptor beta, TEM1/CD248, TEM7R, CLDN6, GPRC5D, CXORF61, CD97, CD179a, ALK, Polysialic acid, PLAC1, GloboH, NY-BR-1, UPK2, HAVCR1, ADRB3, PANX3, GPR20, LY6K, OR51E2, TARP, WT1, NY-ESO-1, LAGE-Ia, MAGE-A1,

legumain, HPV E6, E7, ETV6-AML, sperm protein 17, XAGE1, Tie 2, MAD-CT-1, MAD-CT-2, major histocompatibility complex class I-related gene protein (MR1), urokinase-type plasminogen activator receptor (uPAR), Fos-related antigen 1, p53, p53 mutant, prostein, survivin, telomerase, PCTA-1/Galectin 8, MelanA/MART1, Ras mutant, hTERT, sarcoma translocation breakpoints, ML-IAP, ERG (TMPRSS2 ETS fusion gene), NA17, PAX3, androgen receptor, cyclin B1, MYCN, RhoC, TRP-2, CYP1B I, BORIS, SART3, PAX5, OY-TES1, LCK, AKAP-4, SSX2, RAGE-1, human telomerase reverse transcriptase, RU1, RU2, intestinal carboxyl esterase, mut hsp70-2, CD79a, CD79b, CD72, LAIR1, FCAR, LILRA2, CD300LF, CLEC12A, BST2, EMR2, LY75, GPC3, FCRL5, IGLL1, a neoantigen, CD133, CD15, CD184, CD24, CD56, CD26, CD29, CD44, HLA-A, HLA-B, HLA-C, (HLA-A,B,C) CD49f, CD151 CD340, CD200, trkA, trkB, or trkC, or an antigenic fragment or antigenic portion thereof.

2. ABD targets an antigen characteristic of a T cell

In some embodiments, the antigen binding domain targets an antigen characteristic of a T cell. In some embodiments, the ABD binds an antigen associated with a T cell. In some instances, such an antigen is expressed by a T cell or is located on the surface of a T cell. In some embodiments, the antigen characteristic of a T cell or the T cell associated antigen is selected from a cell surface receptor, a membrane transport protein (e.g., an active or passive transport protein such as, for example, an ion channel protein, a pore-forming protein, etc.), a transmembrane receptor, a membrane enzyme, and/or a cell adhesion protein characteristic of a T cell. In some embodiments, an antigen characteristic of a T cell may be a G protein-coupled receptor, receptor tyrosine kinase, tyrosine kinase associated receptor, receptor-like tyrosine phosphatase, receptor serine/ threonine kinase, receptor guanylyl cyclase, histidine kinase associated receptor, AKT1; AKT2; AKT3; ATF2; BCL10; CALM1; CD3D (CD3 δ); CD3E (CD3 ϵ); CD3G (CD3 γ); CD4; CD8; CD28; CD45; CD80 (B7-1); CD86 (B7-2); CD247 (CD3 ζ); CTLA-4 (CD152); ELK1; ERK1 (MAPK3); ERK2; FOS; FYN; GRAP2 (GADS); GRB2; HLA-DRA; HLA-DRB1; HLA-DRB3; HLA-DRB4; HLA-DRB5; HRAS; IKBKA (CHUK); IKBKB; IKBKE; IKBKG (NEMO); IL2; ITPR1; ITK; JUN; KRAS2; LAT; LCK; MAP2K1 (MEK1); MAP2K2 (MEK2); MAP2K3 (MKK3); MAP2K4 (MKK4); MAP2K6 (MKK6); MAP2K7 (MKK7); MAP3K1 (MEKK1); MAP3K3;

MAP3K4; MAP3K5; MAP3K8; MAP3K14 (NIK); MAPK8 (JNK1); MAPK9 (JNK2); MAPK10 (JNK3); MAPK11 (p38 β); MAPK12 (p38 γ); MAPK13 (p38 δ); MAPK14 (p38 α); NCK; NFAT1; NFAT2; NFKB1; NFKB2; NFKBIA; NRAS; PAK1; PAK2; PAK3; PAK4; PIK3C2B; PIK3C3 (VPS34); PIK3CA; PIK3CB; PIK3CD; PIK3R1; PKCA; PKCB; PKCM; PKCQ; PLCY1; PRF1 (Perforin); PTEN; RAC1; RAF1; RELA; SDF1; SHP2; SLP76; SOS; SRC; TBK1; TCRA; TEC; TRAF6; VAV1; VAV2; or ZAP70.

3. ABD targets an antigen characteristic of an autoimmune or inflammatory disorder

In some embodiments, the antigen binding domain targets an antigen characteristic of an autoimmune or inflammatory disorder. In some embodiments, the ABD binds an antigen associated with an autoimmune or inflammatory disorder. In some instances, the antigen is expressed by a cell associated with an autoimmune or inflammatory disorder. In some embodiments, the autoimmune or inflammatory disorder is selected from chronic graft-vs-host disease (GVHD), lupus, arthritis, immune complex glomerulonephritis, goodpasture syndrome, uveitis, hepatitis, systemic sclerosis or scleroderma, type I diabetes, multiple sclerosis, cold agglutinin disease, Pemphigus vulgaris, Grave's disease, autoimmune hemolytic anemia, Hemophilia A, Primary Sjogren's Syndrome, thrombotic thrombocytopenia purpura, neuromyelitis optica, Evan's syndrome, IgM mediated neuropathy, cryoglobulinemia, dermatomyositis, idiopathic thrombocytopenia, ankylosing spondylitis, bullous pemphigoid, acquired angioedema, chronic urticarial, antiphospholipid demyelinating polyneuropathy, and autoimmune thrombocytopenia or neutropenia or pure red cell aplasias, while exemplary non-limiting examples of alloimmune diseases include allosensitization (see, for example, Blazar et al., 2015, Am. J. Transplant, 15(4):931-41) or xenosensitization from hematopoietic or solid organ transplantation, blood transfusions, pregnancy with fetal allosensitization, neonatal alloimmune thrombocytopenia, hemolytic disease of the newborn, sensitization to foreign antigens such as can occur with replacement of inherited or acquired deficiency disorders treated with enzyme or protein replacement therapy, blood products, and gene therapy. In some embodiments, the antigen characteristic of an autoimmune or inflammatory disorder is selected from a cell surface receptor, an ion channel-

linked receptor, an enzyme-linked receptor, a G protein-coupled receptor, receptor tyrosine kinase, tyrosine kinase associated receptor, receptor-like tyrosine phosphatase, receptor serine/ threonine kinase, receptor guanylyl cyclase, or histidine kinase associated receptor.

In some embodiments, an antigen binding domain of a CAR binds to a ligand expressed on B cells, plasma cells, or plasmablasts. In some embodiments, an antigen binding domain of a CAR binds to CD10, CD19, CD20, CD22, CD24, CD27, CD38, CD45R, CD138, CD319, BAFFR, BCMA, CD28, TNF, interferon receptors, GM-CSF, ZAP-70, LFA-1, CD3 gamma, CD5 or CD2. See, e.g., US 2003/0077249; WO 2017/058753; WO 2017/058850, the contents of which are herein incorporated by reference.

4. ABD targets an antigen characteristic of senescent cells

In some embodiments, the antigen binding domain targets an antigen characteristic of senescent cells, e.g., urokinase-type plasminogen activator receptor (uPAR). In some embodiments, the ABD binds an antigen associated with a senescent cell. In some instances, the antigen is expressed by a senescent cell. In some embodiments, the CAR may be used for treatment or prophylaxis of disorders characterized by the aberrant accumulation of senescent cells, e.g., liver and lung fibrosis, atherosclerosis, diabetes and osteoarthritis.

5. ABD targets an antigen characteristic of an infectious disease

In some embodiments, the antigen binding domain targets an antigen characteristic of an infectious disease. In some embodiments, the ABD binds an antigen associated with an infectious disease. In some instances, the antigen is expressed by a cell affected by an infectious disease. In some embodiments, wherein the infectious disease is selected from HIV, hepatitis B virus, hepatitis C virus, Human herpes virus, Human herpes virus 8 (HHV-8, Kaposi sarcoma-associated herpes virus (KSHV)), Human T-lymphotrophic virus-1 (HTLV-1), Merkel cell polyomavirus (MCV), Simian virus 40 (SV40), Epstein-Barr virus, CMV, human papillomavirus. In some embodiments, the antigen characteristic of an infectious disease is selected from a cell surface receptor, an ion channel-linked receptor, an enzyme-linked receptor, a G protein-coupled receptor, receptor tyrosine kinase, tyrosine kinase

associated receptor, receptor-like tyrosine phosphatase, receptor serine/ threonine kinase, receptor guanylyl cyclase, histidine kinase associated receptor, HIV Env, gp120, or CD4-induced epitope on HIV-1 Env.

6. ABD binds to a cell surface antigen of a cell

In some embodiments, an antigen binding domain binds to a cell surface antigen of a cell. In some embodiments, a cell surface antigen is characteristic of (e.g., expressed by) a particular or specific cell type. In some embodiments, a cell surface antigen is characteristic of more than one type of cell.

In some embodiments, a CAR antigen binding domain binds a cell surface antigen characteristic of a T cell, such as a cell surface antigen on a T cell. In some embodiments, an antigen characteristic of a T cell may be a cell surface receptor, a membrane transport protein (e.g., an active or passive transport protein such as, for example, an ion channel protein, a pore-forming protein, etc.), a transmembrane receptor, a membrane enzyme, and/or a cell adhesion protein characteristic of a T cell. In some embodiments, an antigen characteristic of a T cell may be a G protein-coupled receptor, receptor tyrosine kinase, tyrosine kinase associated receptor, receptor-like tyrosine phosphatase, receptor serine/ threonine kinase, receptor guanylyl cyclase, or histidine kinase associated receptor.

In some embodiments, an antigen binding domain of a CAR binds a T cell receptor. In some embodiments, a T cell receptor may be AKT1; AKT2; AKT3; ATF2; BCL10; CALM1; CD3D (CD3δ); CD3E (CD3ε); CD3G (CD3γ); CD4; CD8; CD28; CD45; CD80 (B7-1); CD86 (B7-2); CD247 (CD3ζ); CTLA-4 (CD152); ELK1; ERK1 (MAPK3); ERK2; FOS; FYN; GRAP2 (GADS); GRB2; HLA-DRA; HLA-DRB1; HLA-DRB3; HLA-DRB4; HLA-DRB5; HRAS; IKBKA (CHUK); IKBKB; IKBKE; IKBKG (NEMO); IL2; ITPR1; ITK; JUN; KRAS2; LAT; LCK; MAP2K1 (MEK1); MAP2K2 (MEK2); MAP2K3 (MKK3); MAP2K4 (MKK4); MAP2K6 (MKK6); MAP2K7 (MKK7); MAP3K1 (MEKK1); MAP3K3; MAP3K4; MAP3K5; MAP3K8; MAP3K14 (NIK); MAPK8 (JNK1); MAPK9 (JNK2); MAPK10 (JNK3); MAPK11 (p38β); MAPK12 (p38γ); MAPK13 (p38δ); MAPK14 (p38α); NCK; NFAT1; NFAT2; NFKB1; NFKB2; NFKBIA; NRAS; PAK1; PAK2; PAK3; PAK4; PIK3C2B; PIK3C3 (VPS34); PIK3CA; PIK3CB; PIK3CD; PIK3R1; PKCA; PKCB; PKCM; PKCQ; PLCY1; PRF1 (Perforin); PTEN;

RAC1; RAF1; RELA; SDF1; SHP2; SLP76; SOS; SRC; TBK1; TCRA; TEC; TRAF6; VAV1; VAV2; or ZAP70.

7. Transmembrane domain

In some embodiments, the CAR transmembrane domain comprises at least a transmembrane region of the alpha, beta or zeta chain of a T cell receptor, CD28, CD3 epsilon, CD45, CD4, CD5, CD8, CD9, CD16, CD22, CD33, CD37, CD64, CD80, CD86, CD134, CD137, CD154, or functional variant thereof. In some embodiments, the transmembrane domain comprises at least a transmembrane region(s) of CD8 α , CD8 β , 4-1BB/CD137, CD28, CD34, CD4, Fc ϵ R1 γ , CD16, OX40/CD134, CD3 ζ , CD3 ϵ , CD3 γ , CD3 δ , TCR α , TCR β , TCR ζ , CD32, CD64, CD64, CD45, CD5, CD9, CD22, CD37, CD80, CD86, CD40, CD40L/CD154, VEGFR2, FAS, and FGFR2B, or functional variant thereof. antigen binding domain binds

8. Signaling domain or plurality of signaling domains

In some embodiments, a CAR described herein comprises one or at least one signaling domain selected from one or more of B7-1/CD80; B7-2/CD86; B7-H1/PD-L1; B7-H2; B7-H3; B7-H4; B7-H6; B7-H7; BTLA/CD272; CD28; CTLA-4; Gi24/VISTA/B7-H5; ICOS/CD278; PD-1; PD-L2/B7-DC; PDCD6); 4-1BB/TNFSF9/CD137; 4-1BB Ligand/TNFSF9; BAFF/BLyS/TNFSF13B; BAFF R/TNFSF13C; CD27/TNFSF7; CD27 Ligand/TNFSF7; CD30/TNFSF8; CD30 Ligand/TNFSF8; CD40/TNFSF5; CD40/TNFSF5; CD40 Ligand/TNFSF5; DR3/TNFSF25; GITR/TNFSF18; GITR Ligand/TNFSF18; HVEM/TNFSF14; LIGHT/TNFSF14; Lymphotoxin-alpha/TNF-beta; OX40/TNFSF4; OX40 Ligand/TNFSF4; RELT/TNFSF19L; TACI/TNFSF13B; TL1A/TNFSF15; TNF-alpha; TNF RII/TNFSF1B); 2B4/CD244/SLAMF4; BLAME/SLAMF8; CD2; CD2F-10/SLAMF9; CD48/SLAMF2; CD58/LFA-3; CD84/SLAMF5; CD229/SLAMF3; CRACC/SLAMF7; NTB-A/SLAMF6; SLAM/CD150); CD2; CD7; CD53; CD82/Kai-1; CD90/Thy1; CD96; CD160; CD200; CD300a/LMIR1; HLA Class I; HLA-DR; Ikaros; Integrin alpha 4/CD49d; Integrin alpha 4 beta 1; Integrin alpha 4 beta 7/LPAM-1; LAG-3; TCL1A; TCL1B; CRTAM; DAP12; Dectin-1/CLEC7A; DPPIV/CD26; EphB6; TIM-1/KIM-1/HAVCR; TIM-4; TSLP; TSLP R; lymphocyte function associated antigen-1 (LFA-1); NKG2C, a CD3 zeta domain, an immunoreceptor tyrosine-based

activation motif (ITAM), CD27, CD28, 4-1BB, CD134/OX40, CD30, CD40, PD-1, ICOS, lymphocyte function-associated antigen-1 (LFA-1), CD2, CD7, LIGHT, NKG2C, B7-H3, a ligand that specifically binds with CD83, or functional fragment thereof.

In some embodiments, the at least one signaling domain comprises a CD3 zeta domain or an immunoreceptor tyrosine-based activation motif (ITAM), or functional variant thereof. In other embodiments, the at least one signaling domain comprises (i) a CD3 zeta domain, or an immunoreceptor tyrosine-based activation motif (ITAM), or functional variant thereof; and (ii) a CD28 domain, or a 4-1BB domain, or functional variant thereof. In yet other embodiments, the at least one signaling domain comprises a (i) a CD3 zeta domain, or an immunoreceptor tyrosine-based activation motif (ITAM), or functional variant thereof; (ii) a CD28 domain or functional variant thereof; and (iii) a 4-1BB domain, or a CD134 domain, or functional variant thereof. In some embodiments, the at least one signaling domain comprises a (i) a CD3 zeta domain, or an immunoreceptor tyrosine-based activation motif (ITAM), or functional variant thereof; (ii) a CD28 domain or functional variant thereof; (iii) a 4-1BB domain, or a CD134 domain, or functional variant thereof; and (iv) a cytokine or costimulatory ligand transgene.

In some embodiments, the at least two signaling domains comprise a CD3 zeta domain or an immunoreceptor tyrosine-based activation motif (ITAM), or functional variant thereof. In other embodiments, the at least two signaling domains comprise (i) a CD3 zeta domain, or an immunoreceptor tyrosine-based activation motif (ITAM), or functional variant thereof; and (ii) a CD28 domain, or a 4-1BB domain, or functional variant thereof. In yet other embodiments, the at least one signaling domain comprises a (i) a CD3 zeta domain, or an immunoreceptor tyrosine-based activation motif (ITAM), or functional variant thereof; (ii) a CD28 domain or functional variant thereof; and (iii) a 4-1BB domain, or a CD134 domain, or functional variant thereof. In some embodiments, the at least two signaling domains comprise a (i) a CD3 zeta domain, or an immunoreceptor tyrosine-based activation motif (ITAM), or functional variant thereof; (ii) a CD28 domain or functional variant thereof; (iii) a 4-1BB domain, or a CD134 domain, or functional variant thereof; and (iv) a cytokine or costimulatory ligand transgene.

In some embodiments, the at least three signaling domains comprise a CD3 zeta domain or an immunoreceptor tyrosine-based activation motif (ITAM), or functional variant thereof. In other embodiments, the at least three signaling domains comprise (i) a CD3 zeta domain, or an immunoreceptor tyrosine-based activation motif (ITAM), or functional variant thereof; and (ii) a CD28 domain, or a 4-1BB domain, or functional variant thereof. In yet other embodiments, the least three signaling domains comprises a (i) a CD3 zeta domain, or an immunoreceptor tyrosine-based activation motif (ITAM), or functional variant thereof; (ii) a CD28 domain or functional variant thereof; and (iii) a 4-1BB domain, or a CD134 domain, or functional variant thereof. In some embodiments, the at least three signaling domains comprise a (i) a CD3 zeta domain, or an immunoreceptor tyrosine-based activation motif (ITAM), or functional variant thereof; (ii) a CD28 domain or functional variant thereof; (iii) a 4-1BB domain, or a CD134 domain, or functional variant thereof; and (iv) a cytokine or costimulatory ligand transgene.

In some embodiments, the CAR comprises a CD3 zeta domain or an immunoreceptor tyrosine-based activation motif (ITAM), or functional variant thereof.

In some embodiments, the CAR comprises (i) a CD3 zeta domain, or an immunoreceptor tyrosine-based activation motif (ITAM), or functional variant thereof; and (ii) a CD28 domain, or a 4-1BB domain, or functional variant thereof.

In some embodiments, the CAR comprises a (i) a CD3 zeta domain, or an immunoreceptor tyrosine-based activation motif (ITAM), or functional variant thereof; (ii) a CD28 domain or functional variant thereof; and (iii) a 4-1BB domain, or a CD134 domain, or functional variant thereof.

In some embodiments, the CAR comprises (i) a CD3 zeta domain, or an immunoreceptor tyrosine-based activation motif (ITAM), or functional variant thereof; (ii) a CD28 domain, or a 4-1BB domain, or functional variant thereof, and/or (iii) a 4-1BB domain, or a CD134 domain, or functional variant thereof.

In some embodiments, the CAR comprises a (i) a CD3 zeta domain, or an immunoreceptor tyrosine-based activation motif (ITAM), or functional variant thereof; (ii) a CD28 domain or functional variant thereof; (iii) a 4-1BB domain, or a CD134 domain, or functional variant thereof; and (iv) a cytokine or costimulatory ligand transgene.

9. Domain which upon successful signaling of the CAR induces expression of a cytokine gene

In some embodiments, a first, second, third, or fourth generation CAR further comprises a domain which upon successful signaling of the CAR induces expression of a cytokine gene. In some embodiments, a cytokine gene is endogenous or exogenous to a target cell comprising a CAR which comprises a domain which upon successful signaling of the CAR induces expression of a cytokine gene. In some embodiments, a cytokine gene encodes a pro-inflammatory cytokine. In some embodiments, a cytokine gene encodes IL-1, IL-2, IL-9, IL-12, IL-18, TNF, or IFN-gamma, or functional fragment thereof. In some embodiments, a domain which upon successful signaling of the CAR induces expression of a cytokine gene is or comprises a transcription factor or functional domain or fragment thereof. In some embodiments, a domain which upon successful signaling of the CAR induces expression of a cytokine gene is or comprises a transcription factor or functional domain or fragment thereof. In some embodiments, a transcription factor or functional domain or fragment thereof is or comprises a nuclear factor of activated T cells (NFAT), an NF-kB, or functional domain or fragment thereof. See, e.g., Zhang, C. et al., Engineering CAR-T cells. Biomarker Research. 5:22 (2017); WO 2016126608; Sha, H. et al. Chimaeric antigen receptor T-cell therapy for tumour immunotherapy. Bioscience Reports Jan 27, 2017, 37 (1).

In some embodiments, the CAR further comprises one or more spacers, e.g., wherein the spacer is a first spacer between the antigen binding domain and the transmembrane domain. In some embodiments, the first spacer includes at least a portion of an immunoglobulin constant region or variant or modified version thereof. In some embodiments, the spacer is a second spacer between the transmembrane domain and a signaling domain. In some embodiments, the second spacer is an oligopeptide, e.g., wherein the oligopeptide comprises glycine and serine residues such as but not limited to glycine-serine doublets. In some embodiments, the CAR comprises two or more spacers, e.g., a spacer between the antigen binding domain and the transmembrane domain and a spacer between the transmembrane domain and a signaling domain.

In some embodiments, any one of the cells described herein comprises a nucleic acid encoding a CAR or a first generation CAR. In some embodiments, a first generation CAR comprises an antigen binding domain, a transmembrane domain, and signaling domain. In some embodiments, a signaling domain mediates downstream signaling during T cell activation.

In some embodiments, the methods and compositions disclosed herein comprise a nucleic acid encoding a CAR or a second generation CAR. In some embodiments, a second generation CAR comprises an antigen binding domain, a transmembrane domain, and two signaling domains. In some embodiments, a signaling domain mediates downstream signaling during T cell activation. In some embodiments, a signaling domain is a costimulatory domain. In some embodiments, a costimulatory domain enhances cytokine production, CAR-T cell proliferation, and/or CAR-T cell persistence during T cell activation.

In some embodiments, any one of the compositions and methods described herein comprises a nucleic acid encoding a CAR or a third generation CAR. In some embodiments, a third generation CAR comprises an antigen binding domain, a transmembrane domain, and at least three signaling domains. In some embodiments, a signaling domain mediates downstream signaling during T cell activation. In some embodiments, a signaling domain is a costimulatory domain. In some embodiments, a costimulatory domain enhances cytokine production, CAR-T cell proliferation, and or CAR-T cell persistence during T cell activation. In some embodiments, a third generation CAR comprises at least two costimulatory domains. In some embodiments, the at least two costimulatory domains are not the same.

In some embodiments, any one of the compositions and methods described herein comprises a nucleic acid encoding a CAR or a fourth generation CAR. In some embodiments, a fourth generation CAR comprises an antigen binding domain, a transmembrane domain, and at least two, three, or four signaling domains. In some embodiments, a signaling domain mediates downstream signaling during T cell activation. In some embodiments, a signaling domain is a costimulatory domain. In some embodiments, a costimulatory domain enhances cytokine production, CAR-T cell proliferation, and or CAR-T cell persistence during T cell activation.

10. ABD comprising an antibody or antigen-binding portion thereof

In some embodiments, a CAR antigen binding domain is or comprises an antibody or antigen-binding portion thereof. In some embodiments, a CAR antigen binding domain is or comprises an scFv or Fab. In some embodiments, a CAR antigen binding domain comprises an scFv or Fab fragment of a CD19 antibody; CD22 antibody; T-cell alpha chain antibody; T-cell β chain antibody; T-cell γ chain antibody; T-cell δ chain antibody; CCR7 antibody; CD3 antibody; CD4 antibody; CD5 antibody; CD7 antibody; CD8 antibody; CD11b antibody; CD11c antibody; CD16 antibody; CD20 antibody; CD21 antibody; CD25 antibody; CD28 antibody; CD34 antibody; CD35 antibody; CD40 antibody; CD45RA antibody; CD45RO antibody; CD52 antibody; CD56 antibody; CD62L antibody; CD68 antibody; CD80 antibody; CD95 antibody; CD117 antibody; CD127 antibody; CD133 antibody; CD137 (4-1 BB) antibody; CD163 antibody; F4/80 antibody; IL-4Ra antibody; Sca-1 antibody; CTLA-4 antibody; GITR antibody; GARP antibody; LAP antibody; granzyme B antibody; LFA-1 antibody; MR1 antibody; uPAR antibody; or transferrin receptor antibody.

In some embodiments, a CAR comprises a signaling domain which is a costimulatory domain. In some embodiments, a CAR comprises a second costimulatory domain. In some embodiments, a CAR comprises at least two costimulatory domains. In some embodiments, a CAR comprises at least three costimulatory domains. In some embodiments, a CAR comprises a costimulatory domain selected from one or more of CD27, CD28, 4-1BB, CD134/OX40, CD30, CD40, PD-1, ICOS, lymphocyte function-associated antigen-1 (LFA-1), CD2, CD7, LIGHT, NKG2C, B7-H3, a ligand that specifically binds with CD83. In some embodiments, if a CAR comprises two or more costimulatory domains, two costimulatory domains are different. In some embodiments, if a CAR comprises two or more costimulatory domains, two costimulatory domains are the same.

In addition to the CARs described herein, various chimeric antigen receptors and nucleotide sequences encoding the same are known in the art and would be suitable for fusosomal delivery and reprogramming of target cells in vivo and in vitro as described herein. See, e.g., WO2013040557; WO2012079000; WO2016030414; Smith T, et al., *Nature Nanotechnology*. 2017. DOI: 10.1038/NNANO.2017.57, the disclosures of which are herein incorporated by reference.

11. Additional Descriptions of CARs

In certain embodiments, the compositions and methods may comprise a polynucleotide encoding a CAR. CARs (also known as chimeric immunoreceptors, chimeric T cell receptors, or artificial T cell receptors) are receptor proteins that have been engineered to give host cells (e.g., T cells) the new ability to target a specific protein. The receptors are chimeric because they combine both antigen-binding and T cell activating functions into a single receptor. The polycistronic vector of the present disclosure may be used to express one or more CARs in a host cell (e.g., a T cell) for use in therapies against various target antigens. The CARs expressed by the one or more expression cassettes may be the same or different. In these embodiments, the CAR may comprise an extracellular binding domain (also referred to as a “binder”) that specifically binds a target antigen, a transmembrane domain, and an intracellular signaling domain. In certain embodiments, the CAR may further comprise one or more additional elements, including one or more signal peptides, one or more extracellular hinge domains, and/or one or more intracellular costimulatory domains. Domains may be directly adjacent to one another, or there may be one or more amino acids linking the domains. The nucleotide sequence encoding a CAR may be derived from a mammalian sequence, for example, a mouse sequence, a primate sequence, a human sequence, or combinations thereof. In the cases where the nucleotide sequence encoding a CAR is non-human, the sequence of the CAR may be humanized. The nucleotide sequence encoding a CAR may also be codon-optimized for expression in a mammalian cell, for example, a human cell. In any of these embodiments, the nucleotide sequence encoding a CAR may be at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to any of the nucleotide sequences disclosed herein. The sequence variations may be due to codon-optimization, humanization, restriction enzyme-based cloning scars, and/or additional amino acid residues linking the functional domains, etc. In certain embodiments, the CAR may comprise a signal peptide at the N-terminus. Non-limiting examples of signal peptides include CD8 α signal peptide, IgK signal peptide, and granulocyte-macrophage colony-stimulating factor receptor subunit alpha (GMCSFR- α , also known as colony stimulating factor 2 receptor subunit alpha

(CSF2RA)) signal peptide, and variants thereof, the amino acid sequences of which are provided in **Table A** below.

Table A. Exemplary sequences of signal peptides

SEQ ID NO:	Sequence	Description
1106	MALPVTALLLPLALLLHAARP	CD8 α signal peptide
1107	METDTLLLWLLLWPGSTG	IgK signal peptide
1108	MLLLVTSLLLCELPHPAFLLIP	GMCSFR- α (CSF2RA) signal peptide
1090	MEFGLSWLFLVAILKGVQCSR	Immunoglobulin heavy chain signal peptide

In certain embodiments, the extracellular binding domain of the CAR may comprise one or more antibodies specific to one target antigen or multiple target antigens. The antibody may be an antibody fragment, for example, an scFv, or a single-domain antibody fragment, for example, a VHH. In certain embodiments, the scFv may comprise a heavy chain variable region (V_H) and a light chain variable region (V_L) of an antibody connected by a linker. The V_H and the V_L may be connected in either order, i.e., V_H-linker-V_L or V_L-linker-V_H. Non-limiting examples of linkers include Whitlow linker, (G₄S)_n (n can be a positive integer, e.g., 1, 2, 3, 4, 5, 6, etc.) linker, and variants thereof. In certain embodiments, the antigen may be an antigen that is exclusively or preferentially expressed on tumor cells, or an antigen that is characteristic of an autoimmune or inflammatory disease. Exemplary target antigens include, but are not limited to, CD5, CD19, CD20, CD22, CD23, CD30, CD70, Kappa, Lambda, and B cell maturation agent (BCMA), G-protein coupled receptor family C group 5 member D (GPRC5D) (associated with leukemias); CS1/SLAMF7, CD38, CD138, GPRC5D, TACI, and BCMA (associated with myelomas); GD2, HER2, EGFR, EGFRvIII, B7H3, PSMA, PSCA, CAIX, CD171, CEA, CSPG4, EPHA2, FAP, FR α , IL-13R α , Mesothelin, MUC1, MUC16, and ROR1 (associated with solid tumors). In any of these embodiments, the extracellular binding domain of the CAR can be codon-optimized for expression in a host cell or have variant sequences to increase functions of the extracellular binding domain.

In certain embodiments, the CAR may comprise a hinge domain, also referred to as a spacer. The terms “hinge” and “spacer” may be used interchangeably in the present disclosure. Non-limiting examples of hinge domains include CD8 α hinge domain, CD28 hinge domain, IgG4 hinge domain, IgG4 hinge-CH2-CH3 domain, and variants thereof, the amino acid sequences of which are provided in **Table B** below.

Table B. Exemplary sequences of hinge domains

SEQ ID NO:	Sequence	Description
1109	TTTPAPRPPTPAPTIASQPLSLRPEACRPAAG GAVHTRGLDFACD	CD8 α hinge domain
1110	IEVMYPPPYLDNEKSNGTIIHVKGKHLCPSP LFGPSKP	CD28 hinge domain
1091	AAAEVMYPPPYLDNEKSNGTIIHVKGKHLCP SPLFGPSKP	CD28 hinge domain
1111	ESKYGPPCPPCP	IgG4 hinge domain
1112	ESKYGPPCPSCP	IgG4 hinge domain
1113	ESKYGPPCPPCPAPEFLGGPSVFLFPPKPKD TLMISRTPEVTCVVVDVSQEDPEVQFNWYVD GVEVHNAKTKPREEQFNSTYRVVSVLTVLHQ DWLNGKEYKCKVSNKGLPSSIEKTISKAKGQ PREPQVYTLPPSQEEMTKNQVSLTCLVKGFY PSDIAVEWESNGQPENNYKTTTPVLDSDGSF FLYSRLTVDKSRWQEGNVFSCSVMHEALHN HYTQKSLSLSLGK	IgG4 hinge-CH2-CH3 domain

In certain embodiments, the transmembrane domain of the CAR may comprise a transmembrane region of the alpha, beta, or zeta chain of a T cell receptor, CD28, CD3 ϵ , CD45, CD4, CD5, CD8, CD9, CD16, CD22, CD33, CD37, CD64, CD80, CD86, CD134, CD137, CD154, or a functional variant thereof, including the human versions of each of these sequences. In other embodiments, the transmembrane domain may comprise a transmembrane region of CD8 α , CD8 β , 4-1BB/CD137, CD28, CD34, CD4, Fc ϵ R1 γ , CD16, OX40/CD134, CD3 ζ , CD3 ϵ , CD3 γ , CD3 δ , TCR α , TCR β , TCR ζ , CD32, CD64, CD64, CD45, CD5, CD9, CD22, CD37, CD80, CD86, CD40, CD40L/CD154, VEGFR2, FAS, and FGFR2B, or a functional variant thereof, including the human versions of each of these sequences. **Table C** provides the amino acid sequences of a few exemplary transmembrane domains.

Table C. Exemplary sequences of transmembrane domains

SEQ ID NO:	Sequence	Description
1114	IYIWAPLAGTCGVLLLSLVITLYC	CD8 α transmembrane domain
1115	FWWLWVVGGLACYSLLVTVAFIIFWW	CD28 transmembrane domain
1214	MFWWLVVVGGLACYSLLVTVAFIIFWW	CD28 transmembrane domain

In certain embodiments, the intracellular signaling domain and/or intracellular costimulatory domain of the CAR may comprise one or more signaling domains selected from B7-1/CD80, B7-2/CD86, B7-H1/PD-L1, B7-H2, B7-H3, B7-H4, B7-H6, B7-H7, BTLA/CD272, CD28, CTLA-4, Gi24/VISTA/B7-H5, ICOS/CD278, PD-1, PD-L2/B7-DC, PDCD6, 4-1BB/TNFSF9/CD137, 4-1BB Ligand/TNFSF9, BAFF/BLyS/TNFSF13B, BAFF R/TNFRSF13C, CD27/TNFRSF7, CD27 Ligand/TNFSF7, CD30/TNFRSF8, CD30 Ligand/TNFSF8, CD40/TNFRSF5, CD40/TNFSF5, CD40 Ligand/TNFSF5, DR3/TNFRSF25, GITR/TNFRSF18, GITR Ligand/TNFSF18, HVEM/TNFRSF14, LIGHT/TNFSF14, Lymphotoxin- α /TNF β , OX40/TNFRSF4, OX40 Ligand/TNFSF4, RELT/TNFRSF19L, TACI/TNFRSF13B, TL1A/TNFSF15, TNF α , TNF RII/TNFRSF1B, 2B4/CD244/SLAMF4, BLAME/SLAMF8, CD2, CD2F-10/SLAMF9, CD48/SLAMF2, CD58/LFA-3, CD84/SLAMF5, CD229/SLAMF3, CRACC/SLAMF7, NTB-A/SLAMF6, SLAM/CD150, CD2, CD7, CD53, CD82/Kai-1, CD90/Thy1, CD96, CD160, CD200, CD300a/LMIR1, HLA Class I, HLA-DR, Ikaros, Integrin α 4/CD49d, Integrin α 4 β 1, Integrin α 4 β 7/LPAM-1, LAG-3, TCL1A, TCL1B, CRTAM, DAP12, Dectin-1/CLEC7A, DPPIV/CD26, EphB6, TIM-1/KIM-1/HAVCR, TIM-4, TSLP, TSLP R, lymphocyte function associated antigen-1 (LFA-1), NKG2C, CD3 ζ , an immunoreceptor tyrosine-based activation motif (ITAM), CD27, CD28, 4-1BB, CD134/OX40, CD30, CD40, PD-1, ICOS, lymphocyte function-associated antigen-1 (LFA-1), CD2, CD7, LIGHT, NKG2C, B7-H3, a ligand that specifically binds with CD83, and a functional variant thereof including the human versions of each of these sequences. In some embodiments, the intracellular signaling domain and/or intracellular costimulatory domain comprises one or more signaling domains

selected from a CD3 ζ domain, an ITAM, a CD28 domain, 4-1BB domain, or a functional variant thereof. **Table D** provides the amino acid sequences of a few exemplary intracellular costimulatory and/or signaling domains. In certain embodiments, as in the case of tisagenlecleucel as described below, the CD3 ζ signaling domain of SEQ ID NO:1118 may have a mutation, *e.g.*, a glutamine (Q) to lysine (K) mutation, at amino acid position 14 (see SEQ ID NO:1215).

Table D. Exemplary sequences of intracellular costimulatory and/or signaling domains

SEQ ID NO:	Sequence	Description
1116	KRGRKKLLYIFKQPFMRPVQTTQEEDGC SCRFPEEEEEGGCEL	4-1BB costimulatory domain
1117	RSKRSRLLHSDYMNMTPRRPGPTRKHY QPYAPPRDFAAYRS	CD28 costimulatory domain
1091	RSKRSRGGHSDYMNMTPRRPGPTRKH YQPYAPPRDFAAYRS	CD28 costimulatory domain (LL>GG mutant)
1118	RVKFSRSADAPAYQQGQNQLYNELNLG RREEYDVLDKRRGRDPEMGGKPRRKN PQEGLYNELQKDKMAEAYSEIGMKGER RRGKGHDGLYQGLSTATKDTYDALHMQ ALPPR	CD3 ζ signaling domain
1215	RVKFSRSADAPAYKQGQNQLYNELNLG RREEYDVLDKRRGRDPEMGGKPRRKN PQEGLYNELQKDKMAEAYSEIGMKGER RRGKGHDGLYQGLSTATKDTYDALHMQ ALPPR	CD3 ζ signaling domain (with Q to K mutation at position 14)

In certain embodiments where the polycistronic vector encodes two or more CARs, the two or more CARs may comprise the same functional domains, or one or more different functional domains, as described. For example, the two or more CARs may comprise different signal peptides, extracellular binding domains, hinge domains, transmembrane domains, costimulatory domains, and/or intracellular signaling domains, in order to minimize the risk of recombination due to sequence similarities. Or, alternatively, the two or more CARs may comprise the same domains. In the cases where the same domain(s) and/or backbone are used, it is optional to

introduce codon divergence at the nucleotide sequence level to minimize the risk of recombination.

CD19 CAR

In some embodiments, the CAR is a CD19 CAR ("CD19-CAR"), and in these embodiments, the polycistronic vector comprises an expression cassette that contains a nucleotide sequence encoding a CD19 CAR. In some embodiments, the CD19 CAR may comprise a signal peptide, an extracellular binding domain that specifically binds CD19, a hinge domain, a transmembrane domain, an intracellular costimulatory domain, and/or an intracellular signaling domain in tandem.

In some embodiments, the signal peptide of the CD19 CAR comprises a CD8 α signal peptide. In some embodiments, the CD8 α signal peptide comprises or consists of an amino acid sequence set forth in SEQ ID NO:1106 or an amino acid sequence that is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in of SEQ ID NO:1106. In some embodiments, the signal peptide comprises an IgK signal peptide. In some embodiments, the IgK signal peptide comprises or consists of an amino acid sequence set forth in SEQ ID NO:1107 or an amino acid sequence that is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in of SEQ ID NO:1107. In some embodiments, the signal peptide comprises a GMCSFR- α or CSF2RA signal peptide. In some embodiments, the GMCSFR- α or CSF2RA signal peptide comprises or consists of an amino acid sequence set forth in SEQ ID NO:1108 or an amino acid sequence that is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in of SEQ ID NO:1108.

In some embodiments, the extracellular binding domain of the CD19 CAR is specific to CD19, for example, human CD19. The extracellular binding domain of the CD19 CAR can be codon-optimized for expression in a host cell or to have variant sequences to increase functions of the extracellular binding domain. In some embodiments, the extracellular binding domain comprises an immunogenically active portion of an immunoglobulin molecule, for example, an scFv.

In some embodiments, the extracellular binding domain of the CD19 CAR comprises an scFv derived from the FMC63 monoclonal antibody (FMC63), which comprises the heavy chain variable region (V_H) and the light chain variable region (V_L) of FMC63 connected by a linker. FMC63 and the derived scFv have been described in Nicholson et al., Mol. Immun. 34(16-17):1157-1165 (1997) and PCT Application Publication No. WO2018/213337, the entire contents of each of which are incorporated by reference herein. In some embodiments, the amino acid sequences of the entire FMC63-derived scFv (also referred to as FMC63 scFv) and its different portions are provided in **Table E** below. In some embodiments, the CD19-specific scFv comprises or consists of an amino acid sequence set forth in SEQ ID NO:1075, 1119, 1120, or 1125, or an amino acid sequence that is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in SEQ ID NO:1075, 1119, 1120, or 1125. In some embodiments, the CD19-specific scFv may comprise one or more CDRs having amino acid sequences set forth in SEQ ID NOs: 1121-1123 and 1126-1128. In some embodiments, the CD19-specific scFv may comprise a light chain with one or more CDRs having amino acid sequences set forth in SEQ ID NOs: 1121-1123. In some embodiments, the CD19-specific scFv may comprise a heavy chain with one or more CDRs having amino acid sequences set forth in SEQ ID NOs: 1126-1128. In any of these embodiments, the CD19-specific scFv may comprise one or more CDRs comprising one or more amino acid substitutions, or comprising a sequence that is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical), to any of the sequences identified. In some embodiments, the extracellular binding domain of the CD19 CAR comprises or consists of the one or more CDRs as described herein.

In some embodiments, the linker linking the V_H and the V_L portions of the scFv is a Whitlow linker having an amino acid sequence set forth in SEQ ID NO:1124. In some embodiments, the Whitlow linker may be replaced by a different linker, for example, a 3xG4S linker having an amino acid sequence set forth in SEQ ID NO:1130, which gives rise to a different FMC63-derived scFv having an amino acid sequence set forth in SEQ ID NO:1129. In certain of these embodiments, the CD19-specific scFv comprises or consists of an amino acid sequence set forth in SEQ ID

NO:1129 or an amino acid sequence that is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in of SEQ ID NO:1129.

Table E. Exemplary sequences of anti-CD19 scFv and components

SEQ ID NO:	Amino Acid Sequence	Description
1119	DIQMTQTTSSLSASLGDRVTISCRA SQDISKYLNWYQQKPDGTVKLLIYH TSRLHSGVPSRFSGSGSGTDYSLTI SNLEQEDIATYFCQQGNTLPYTFGG GTKLEITGSTSGSGKPGSGEGSTK GEVKLQESGPGLVAPSQSLSVTCT VSGVSLPDYGVSWIRQPPRKGLEW LGVIWGSETTYNSALKSRLTIKDN SKSQVFLKMNSLQTDDTAIYYCAKH YYYGGSYAMDYWGQGTSVTVSS	Anti-CD19 FMC63 scFv entire sequence, with Whitlow linker
1120	DIQMTQTTSSLSASLGDRVTISCRA SQDISKYLNWYQQKPDGTVKLLIYH TSRLHSGVPSRFSGSGSGTDYSLTI SNLEQEDIATYFCQQGNTLPYTFGG GTKLEIT	Anti-CD19 FMC63 scFv light chain variable region
1121	QDISKY	Anti-CD19 FMC63 scFv light chain CDR1
1122	HTS	Anti-CD19 FMC63 scFv light chain CDR2
1123	QQGNTLPYT	Anti-CD19 FMC63 scFv light chain CDR3
1124	GSTSGSGKPGSGEGSTKG	Whitlow linker
1125	EVKLQESGPGLVAPSQSLSVTCTV SGVSLPDYGVSWIRQPPRKGLEWL GVIWGSETTYNSALKSRLTIKDNS KSQVFLKMNSLQTDDTAIYYCAKHY YYGGSYAMDYWGQGTSVTVSS	Anti-CD19 FMC63 scFv heavy chain variable region
1126	GVSLPDYG	Anti-CD19 FMC63 scFv heavy chain CDR1
1127	IWGSETT	Anti-CD19 FMC63 scFv heavy chain CDR2
1128	AKHYYYGGSYAMDY	Anti-CD19 FMC63 scFv heavy chain CDR3

SEQ ID NO:	Amino Acid Sequence	Description
1129	DIQMTQTTSSLSASLGDRVTISCRA SQDISKYLNWYQQKPDGTVKLLIYH TSRLHSGVPSRFSGSGSGTDYSLTI SNLEQEDIATYFCQQGNTLPYTFGG GTKLEITGGGGSGGGGSGGGGSE VKLQESGPGLVAPSQSLSVTCTVS GVSLPDYGVSWIRQPPRKGLEWLG VIWGSETTYNSALKSRLTIKDNSK SQVFLKMNSLQTDDEAIYYCAKHYY YGGSYAMDYWGQGTSTVTVSS	Anti-CD19 FMC63 scFv entire sequence, with 3xG4S linker (SEQ ID NO: 1130)
1130	GGGGSGGGGSGGGGS	3xG4S linker

In some embodiments, the extracellular binding domain of the CD19 CAR is derived from an antibody specific to CD19, including, for example, SJ25C1 (Bejcek et al., Cancer Res. 55:2346-2351 (1995)), HD37 (Pezutto et al., J. Immunol. 138(9):2793-2799 (1987)), 4G7 (Meeker et al., Hybridoma 3:305-320 (1984)), B43 (Bejcek (1995)), BLY3 (Bejcek (1995)), B4 (Freedman et al., 70:418-427 (1987)), B4 HB12b (Kansas & Tedder, J. Immunol. 147:4094-4102 (1991); Yazawa et al., Proc. Natl. Acad. Sci. USA 102:15178-15183 (2005); Herbst et al., J. Pharmacol. Exp. Ther. 335:213-222 (2010)), BU12 (Callard et al., J. Immunology, 148(10): 2983-2987 (1992)), and CLB-CD19 (De Rie Cell. Immunol. 118:368-381(1989)). In any of these embodiments, the extracellular binding domain of the CD19 CAR can comprise or consist of the V_H, the V_L, and/or one or more CDRs of any of the antibodies.

In some embodiments, the hinge domain of the CD19 CAR comprises a CD8 α hinge domain, for example, a human CD8 α hinge domain. In some embodiments, the CD8 α hinge domain comprises or consists of an amino acid sequence set forth in SEQ ID NO:1109 or an amino acid sequence that is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in of SEQ ID NO:1109. In some embodiments, the hinge domain comprises a CD28 hinge domain, for example, a human CD28 hinge domain. In some embodiments, the CD28 hinge domain comprises or consists of an amino acid sequence set forth in SEQ ID NO:1110 or an amino acid sequence that is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in of

SEQ ID NO:1110. In some embodiments, the hinge domain comprises an IgG4 hinge domain, for example, a human IgG4 hinge domain. In some embodiments, the IgG4 hinge domain comprises or consists of an amino acid sequence set forth in SEQ ID NO:1111 or SEQ ID NO:1112, or an amino acid sequence that is at least 80% identical (*e.g.*, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in of SEQ ID NO: 1111 or SEQ ID NO: 1112. In some embodiments, the hinge domain comprises a IgG4 hinge-Ch2-Ch3 domain, for example, a human IgG4 hinge-Ch2-Ch3 domain. In some embodiments, the IgG4 hinge-Ch2-Ch3 domain comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1113 or an amino acid sequence that is at least 80% identical (*e.g.*, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in of SEQ ID NO: 1113.

In some embodiments, the transmembrane domain of the CD19 CAR comprises a CD8 α transmembrane domain, for example, a human CD8 α transmembrane domain. In some embodiments, the CD8 α transmembrane domain comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1114 or an amino acid sequence that is at least 80% identical (*e.g.*, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in SEQ ID NO: 1114. In some embodiments, the transmembrane domain comprises a CD28 transmembrane domain, for example, a human CD28 transmembrane domain. In some embodiments, the CD28 transmembrane domain comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1115 or an amino acid sequence that is at least 80% identical (*e.g.*, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in SEQ ID NO: 1115.

In some embodiments, the intracellular costimulatory domain of the CD19 CAR comprises a 4-1BB costimulatory domain. 4-1BB, also known as CD137, transmits a potent costimulatory signal to T cells, promoting differentiation and enhancing long-term survival of T lymphocytes. In some embodiments, the 4-1BB costimulatory domain is human. In some embodiments, the 4-1BB costimulatory domain

comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1116 or an amino acid sequence that is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in SEQ ID NO: 1116. In some embodiments, the intracellular costimulatory domain comprises a CD28 costimulatory domain. CD28 is another co-stimulatory molecule on T cells. In some embodiments, the CD28 costimulatory domain is human. In some embodiments, the CD28 costimulatory domain comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1117 or an amino acid sequence that is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in SEQ ID NO: 1117. In some embodiments, the intracellular costimulatory domain of the CD19 CAR comprises a 4-1BB costimulatory domain and a CD28 costimulatory domain as described.

In some embodiments, the intracellular signaling domain of the CD19 CAR comprises a CD3 zeta (ζ) signaling domain. CD3 ζ associates with T cell receptors (TCRs) to produce a signal and contains immunoreceptor tyrosine-based activation motifs (ITAMs). The CD3 ζ signaling domain refers to amino acid residues from the cytoplasmic domain of the zeta chain that are sufficient to functionally transmit an initial signal necessary for T cell activation. In some embodiments, the CD3 ζ signaling domain is human. In some embodiments, the CD3 ζ signaling domain comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1118 or an amino acid sequence that is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in SEQ ID NO: 1118.

In some embodiments, the polycistronic vector comprises an expression cassette that contains a nucleotide sequence encoding a CD19 CAR, including, for example, a CD19 CAR comprising the CD19-specific scFv having sequences set forth in SEQ ID NO: 1119 or SEQ ID NO: 1129, the CD8 α hinge domain of SEQ ID NO: 1109, the CD8 α transmembrane domain of SEQ ID NO: 1114, the 4-1BB costimulatory domain of SEQ ID NO: 1116, the CD3 ζ signaling domain of SEQ ID NO: 1118, and/or variants (i.e., having a sequence that is at least 80% identical, for example, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least

98%, or at least 99 identical to the disclosed sequence) thereof. In any of these embodiments, the CD19 CAR may additionally comprise a signal peptide (e.g., a CD8 α signal peptide) as described.

In some embodiments, the polycistronic vector comprises an expression cassette that contains a nucleotide sequence encoding a CD19 CAR, including, for example, a CD19 CAR comprising the CD19-specific scFv having sequences set forth in SEQ ID NO: 1119 or SEQ ID NO: 1129, the IgG4 hinge domain of SEQ ID NO: 1111 or SEQ ID NO: 1112, the CD28 transmembrane domain of SEQ ID NO: 1115, the 4-1BB costimulatory domain of SEQ ID NO: 1116, the CD3 ζ signaling domain of SEQ ID NO: 1118, and/or variants (i.e., having a sequence that is at least 80% identical, for example, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99 identical to the disclosed sequence) thereof. In any of these embodiments, the CD19 CAR may additionally comprise a signal peptide (e.g., a CD8 α signal peptide) as described.

In some embodiments, the polycistronic vector comprises an expression cassette that contains a nucleotide sequence encoding a CD19 CAR, including, for example, a CD19 CAR comprising the CD19-specific scFv having sequences set forth in SEQ ID NO: 1075, SEQ ID NO: 1119, or SEQ ID NO: 1129, the CD28 hinge domain of SEQ ID NO: 1110, the CD28 transmembrane domain of SEQ ID NO: 1115, the CD28 costimulatory domain of SEQ ID NO: 1117, the CD3 ζ signaling domain of SEQ ID NO: 1118, and/or variants (i.e., having a sequence that is at least 80% identical, for example, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99 identical to the disclosed sequence) thereof. In any of these embodiments, the CD19 CAR may additionally comprise a signal peptide (e.g., a CD8 α signal peptide) as described.

In some embodiments, the polycistronic vector comprises an expression cassette that contains a nucleotide sequence encoding a CD19 CAR comprising the CD19-specific scFv having the sequence set forth in SEQ ID NO: 1075 (see **Table 14**), the CD8 hinge domain of SEQ ID NO: 1109, the CD8 transmembrane domain of SEQ ID NO: 1114, the 4-1BB costimulatory domain of SEQ ID NO: 1116, the CD3 ζ signaling domain of SEQ ID NO: 1118, and/or variants (i.e., having a sequence that is at least 80% identical, for example, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99 identical to the disclosed

sequence) thereof. In any of these embodiments, the CD19 CAR may additionally comprise a signal peptide (e.g., a CD8 α signal peptide) as described.

In some embodiments, the polycistronic vector comprises an expression cassette that contains a nucleotide sequence encoding a CD19 CAR as set forth in SEQ ID NO: 1216 or is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the nucleotide sequence set forth in SEQ ID NO: 1216 (see **Table F**).

The encoded CD19 CAR has a corresponding amino acid sequence set forth in SEQ ID NO: 1217 or is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in of SEQ ID NO: 1217, with the following components: CD8 α signal peptide, FMC63 scFv (V_L-Whitlow linker-V_H), CD8 α hinge domain, CD8 α transmembrane domain, 4-1BB costimulatory domain, and CD3 ζ signaling domain.

In some embodiments, the polycistronic vector comprises an expression cassette that contains a nucleotide sequence encoding a commercially available embodiment of CD19 CAR. Non-limiting examples of commercially available embodiments of CD19 CARs expressed and/or encoded by T cells include tisagenlecleucel, lisocabtagene maraleucel, axicabtagene ciloleucel, and brexucabtagene autoleucel.

In some embodiments, the polycistronic vector comprises an expression cassette that contains a nucleotide sequence encoding tisagenlecleucel or portions thereof. Tisagenlecleucel comprises a CD19 CAR with the following components: CD8 α signal peptide, FMC63 scFv (V_L-3xG4S linker-V_H), CD8 α hinge domain, CD8 α transmembrane domain, 4-1BB costimulatory domain, and CD3 ζ signaling domain. The nucleotide and amino acid sequence of the CD19 CAR in tisagenlecleucel are provided in **Table F**, with annotations of the sequences provided in **Table G**.

In some embodiments, the polycistronic vector comprises an expression cassette that contains a nucleotide sequence encoding lisocabtagene maraleucel or portions thereof. Lisocabtagene maraleucel comprises a CD19 CAR with the following components: GMCSFR- α or CSF2RA signal peptide, FMC63 scFv (V_L-Whitlow linker-V_H), IgG4 hinge domain, CD28 transmembrane domain, 4-1BB costimulatory domain, and CD3 ζ signaling domain. The nucleotide and amino acid sequence of

the CD19 CAR in lisocabtagene maraleucel are provided in **Table F**, with annotations of the sequences provided in **Table H**.

In some embodiments, the polycistronic vector comprises an expression cassette that contains a nucleotide sequence encoding axicabtagene ciloleucel or portions thereof. Axicabtagene ciloleucel comprises a CD19 CAR with the following components: GMCSFR- α or CSF2RA signal peptide, FMC63 scFv (V_L-Whitlow linker-V_H), CD28 hinge domain, CD28 transmembrane domain, CD28 costimulatory domain, and CD3 ζ signaling domain. The nucleotide and amino acid sequence of the CD19 CAR in axicabtagene ciloleucel are provided in **Table F**, with annotations of the sequences provided in **Table I**.

In some embodiments, the polycistronic vector comprises an expression cassette that contains a nucleotide sequence encoding brexucabtagene autoleucel or portions thereof. Brexucabtagene autoleucel comprises a CD19 CAR with the following components: GMCSFR- α signal peptide, FMC63 scFv, CD28 hinge domain, CD28 transmembrane domain, CD28 costimulatory domain, and CD3 ζ signaling domain. In some embodiments, the polycistronic vector comprises an expression cassette that contains a nucleotide sequence encoding a CD19 CAR as set forth in SEQ ID NO: 1131, 1133, or 1135, or is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the nucleotide sequence set forth in SEQ ID NO: 1131, 1133, or 1135. The encoded CD19 CAR has a corresponding amino acid sequence set forth in SEQ ID NO: 1132, 1134, or 1136, respectively, or is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in of SEQ ID NO: 1132, 1134, or 1136, respectively.

Table F. Exemplary sequences of CD19 CARs

SEQ ID NO:	Sequence	Description
1216	atggccttaccagtgaccgccttgctcctgccgctggcctg ctgctccacgcgcgagccaggccggacatccagatgacacag actacatcctccctgtctgcctctctgggagacagagtcacc atcagttgcagggcaagtcaggacattagtaaattttaaa ttggtatcagcagaaaccagatggaactgttaaactcctga tctaccatacatcaagattacactcaggagtcctcatcaagg ttcagtggcagtggtctggaacagattattctctcaccatta gcaacctggagcaagaagatattgccactacttttgccaa	Exemplary CD19 CAR nucleotide sequence

SEQ ID NO:	Sequence	Description
	cagggttaatacgttccgtacacgttcggaggggggacc aagctggagatcacaggctccacctctggatccggcaag cccggatctggcgagggatccaccaagggcgaggtgaa actgcaggagtcaggacctggcctggtggcgccctcac gagcctgtccgtcacatgcactgtctcaggggtctcattacc cgactatggtgtaagctggattcgccagcctccacgaaag ggtctggagtggctgggagtaatatgggtagtgaaacca catactataattcagctctcaaatccagactgaccatcatca aggacaactccaagagccaagttttctaaaaatgaacag tctgcaaactgatgacacagccatttactactgtgccaaac attattactacggtgtagctatgctatggactactggggcc aaggaacctcagtcaccgtctcctcaaccacgacgccag cgccgagaccaccaacaccggcgcccaccatcgcgctcg cagcccctgtccctgcgcccagaggcgtgccggccagcg gcgggggggcgagtgacacgagggggtggacttcgc ctgtgatatctacatctgggcgcccctggccgggactgtgg ggtccttctcctgtcactggttatcaccccttactgcaaacgg ggcagaaaagaaactcctgtatatattcaacaaccatttat gagaccagtacaaaactactcaagaggaagatggctgtag ctgccgatttcagaagaagaagaaggaggatgtgaact gagagtgaagttcagcaggagcgcagacgcccccgct accagcagggccagaaccagctctataacgagctcaatc taggacgaagagaggagtacgatgttttgacaagagac gtggccgggaccctgagatggggggaaagccgagaag gaagaaccctcaggaaggcctgtacaatgaactgcaga aagataagatggcggaggcctacagtgagattgggatga aaggcgagcgccggagggggcaaggggcacgatggcct ttaccagggtctcagtacagccaccaaggacacctacga cgcccttcacatgcaggccctgccccctcgc	
1217	MALPVTALLPLALLHAARPDIQMTQTTS SLSASLGDRVTISCRASQDISKYLNWYQQ KPDGTVKLLIYHTSRLHSGVPSRFSGSGS GTDYSLTISNLEQEDIATYFCQQGNTLPYT FGGGTKLEITGSTSGSGKPGSGEGSTKGE VKLQESGPGLVAPSQSLSVTCTVSGVSLP DYGVSWIRQPPRKGLEWLGVWIGSETTY YNSALKSRLTIKDNSKSQVFLKMNSLQTD DTAIYYCAKHYYYGGSYAMDYWGQGTSTV TVSSTTTTPAPRPPTPAPTIASQPLSLRPEA CRPAAGGAVHTRGLDFACDIYWAPLAGT CGVLLLSLVITLYCKRGRKKLLYIFKQPFM RPVQTTQEEDGCSCRFPEEEEGGCELRV KFSRSADAPAYQQGQNQLYNELNLGRRE EYDVLDKRRGRDPEMGGKPRRKNPQEGL YNELQKDKMAEAYSEIGMKGERRRGKGH DGLYQGLSTATKDTYDALHMQALPPR	Exemplary CD19 CAR amino acid sequence

SEQ ID NO:	Sequence	Description
1231	DIQMTQTTSSLSASLGDRVTISCRASQDIS KYLNWYQQKPDGTVKLLIYHTSRLHSGVP SRFSGSGSGTDYSLTISNLEQEDIATYFCQ QGNTLPYTFGGGKLEITGSTSGSGKPGS GEGSTKGEVKLQESGPGLVAPSQSLSVT CTVSGVSLPDYGVSWIRQPPRKGLEWLG VIWGSETTYNSALKSRLTIKDNSKSQVFL KMNSLQTDDTAIYYCAKHYYYGGSYAMD YWGQGSVTVSSTTTTPAPRPPTPAPTAS QPLSLRPEACRPAAGGAVHTRGLDFACDI YIWAPLAGTCGVLLLSLVITLYCKRGRKKL LYIFKQPFMRPVQTTQEEDGCSCRFPEEE EGGCELRVKFSRSADAPAYQQGQNQLYN ELNLGRREEYDVLDKRRGRDPEMGGKPR RKNPQEGLYNELQKDKMAEAYSEIGMKG ERRRGKHDGLYQGLSTATKDTYDALHM QALPPR	Exemplary CD19 CAR amino acid sequence (no signal sequence)
1131	atggccttaccagtgaccgccttgctcctgccgctggccttg ctgctccacgccgccaggccggacatccagatgacacag actacatcctccctgtctgcctctctgggagacagagtcacc atcagttgcagggaagtcaggacattagtaaatatttaa ttggtatcagcagaaaccagatggaactgttaaactcctga tctaccatacatcaagattacactcaggagtcctcatcaagg ttcagtggcagtggtctggaacagattattctctcaccatta gcaacctggagcaagaagatattgccactacttttgcaa cagggtataacgcttcggtacacgttcggaggggggacc aagctggagatcacagggtggcgggtggtcggcggtggt gggtcgggtggcgcggtatctgaggtgaaactgcaggag tcaggacctggcctggtggcgccctcacagagcctgtccg tcacatgcactgtctcaggggtctcattaccgactatggtgt aagctggattcgccagcctccacgaaagggtctggagt gctgggagtaatatggggtagtgaaccacatactataatt cagctctcaaatccagactgacctcatcaaggacaactc caagagccaagttttctaaaaatgaacagtctgcaaactg atgacacagccatttactactgtgcaaacattattactacg gtggtagctatgctatggactactggggccaaggaacctc agtcaccgtctcctcaaccacgacgccagcgccgacc accaacaccggcgcccaccatcgctcgcagcccctgtc cctgcgcccagaggcgtgccggccagcgcgggggggc gcagtgcacacgaggggggtggacttcgcctgtgatct acatctggcgcccttgccgggactgtgggtccttctcc tgtactgggtatcacccttactgcaaacggggcagaaaag aaactcctgtatatattcaaacaaccattatgagaccagta caaactactcaagaggaagatggctgtagctgccgatttc cagaagaagaagaaggaggatgtgaactgagagtga gttcagcaggagcgacagcggcggtacaagcagg gccagaaccagctctataacgagctcaatctaggacgaa	Tisagenlecleucel CD19 CAR nucleotide sequence

SEQ ID NO:	Sequence	Description
	gagaggagtacgatgttttgacaagagacgtggccggg accctgagatgggggaaagccgagaaggaagaacc tcaggaaggcctgtacaatgaactgcagaaagataagat ggcggaggcctacagtgagattgggatgaaaggcgagc gccggaggggcaaggggcacgatggccttaccagggt ctcagtacagccaccaaggacacctacgacgccttcac atgcaggccctgccccctcgc	
1132	MALPVTALLPLALLLHAARPDIQMTQTTS SLSASLGDRVTISCRASQDISKYLNWYQQ KPDGTVKLLIYHTSRLHSGVPSRFSGSGS GTDYSLTISNLEQEDIATYFCQQGNTLPYT FGGGTKLEITGGGSGGGGSGGGGSEVK LQESGPGLVAPSQSLSVTCTVSGVSLPDY GVSWIRQPPRKGLEWLGVWGETTYYN SALKSRLTIKDNSKSQVFLKMNSLQDDT AIYYCAKHYYYGGSYAMDYWGQGTSTTV SSTTTPAPRPPTPAPTIASQPLSLRPEACR PAAGGAVHTRGLDFACDIYIWAPLAGTCG VLLLSLVITLYCKRGRKKLLYIFKQPFMRPV QTTQEEDGCSCRFPEEEEEGGCELRVKFS RSADAPAYKQGQNQLYNELNLGRREEYD VLDKRRGRDPEMGGKPRRKNPQEGLYNE LQKDKMAEAYSEIGMKGERRRGKGHDL YQGLSTATKDTYDALHMQALPPR	Tisagenlecleucel CD19 CAR amino acid sequence
1133	atgctgctgctggtgaccagcctgctgctgtgcgagctgcc ccaccccgcttctgctgatccccgacatccagatgaccc agaccacctccagcctgagcgccagcctggcgaccgg gtgaccatcagctgccgggcccagccaggacatcagcaa gtacctgaactggtatcagcagaagcccgacggcaccgt caagctgctgatctaccacaccagccggctgcacagcgg cgtgccagccggttagcggcagcggctccggcaccga ctacagcctgaccatctccaacctggaacaggaagatc gccacctaactttgccagcagggaacacactgccctaca ccttggcgggcggaacaaagctggaaatcacccgcagca cctccggcagcggcaagcctggcagcggcgagggcag caccaagggcgaggtgaagctgcaggaaagcggccct ggcctggtggccccagccagagcctgagcgtgacctgc accgtgagcggcgtgagcctgcccgactacggcgtgagc tgatccggcagccccccaggaaggcctggaatggctg ggcgtgatctgggagcagcagaccactactacaacagc gccctgaagagccggctgaccatcatcaaggacaacag caagagccaggtgttctgaagatgaacagcctgcagac cgacgacaccgccatctactactgcgccaagcactactac tacggcggcagctacgccatggactactggggccagggc accagcgtgaccgtgagcagcgaatctaagtacggaccg ccctgcccccttgccctatgttctgggtgctggtggtgctg	Lisocabtagene maraleucel CD19 CAR nucleotide sequence

SEQ ID NO:	Sequence	Description
	gaggcgtgctggcctgctacagcctgctggcaccgtggc cttcatcatcttttgggtgaaacggggcagaaagaaactcc tgtatatattcaaacaaccatttatgagaccagtacaaacta ctcaagaggaagatggctgtagctgccgatttccagaaga agaagaaggaggatgtgaactgcgggtgaagttcagca gaagcgccgacgcccctgcctaccagcagggccagaat cagctgtacaacgagctgaacctgggcagaagggaaga gtacgacgtcctggataagcggagaggccgggaccctg agatgggcggaagcctcggcggaagaacccccagga aggcctgtataacgaactgcagaaagacaagatggccg aggcctacagcgagatcgcatgaaggcgagcggag gcggggcaagggccacgacggcctgtatcagggcctgtc caccgccaccaaggatacctacgacgccctgcacatgca ggcctgcccccaagg	
1134	MLLLVTSLLLCELPHPAFLLIPDIQMTQTTS SLSASLGDRVTISCRASQDISKYLNWYQQ KPDGTVKLLIYHTSRLHSGVPSRFSGSGS GTDYSLTISNLEQEDIATYFCQQGNTLPYT FGGGTKLEITGSTSGSGKPGSGEGSTKGE VKLQESGPGLVAPSQSLSVTCTVSGVSLP DYGVSWIRQPPRKGLEWLGVWGETTY YNSALKSRLTIKDNSKSQVFLKMNSLQTD DTAIYYCAKHYYYGGSYAMDYWGQTSV TVSSESKYGPPCPPCPMFWLVVVGVL ACYSLLVTVAFIIFWVKRGRKLLYIFKQPF MRPVQTTQEEDGCSCRFPEEEEEGGCEL VKFSRSADAPAYQQGQNQLYNELNLGRR EEYDVLDKRRGRDPEMGGKPRRKNPQE GLYNELQKDKMAEAYSEIGMKGERRRGK GHDGLYQGLSTATKDTYDALHMQALPPR	Lisocabtagene maraleucel CD19 CAR amino acid sequence
1135	atgcttctcctggtgacaagccttctgctctgtgagttaccac accagcattcctcctgatcccagacatccagatgacaca gactacatcctcctgtctgcctctctgggagacagagtca ccatcagttgcagggcaagtcaggacattagtaaataatta aattggtatcagcagaaaccagatggaactgttaaactcct gatctaccatacatcaagattacactcaggagtcctcatcaa ggttcagtggcagtggtctggaacagattattctctacca ttagcaacctggagcaagaagatattgccactacttttgc aacagggtaatacgttccgtacacgttcggagggggga ctaagttgaaataacaggctccacctctggatccggcaa gcccggatctggcaggggatccaccaagggcgaggtga aactgcaggagtcaggacctggcctggtggcgccctcac agagcctgtccgtcacatgcactgtctcaggggtctcattac ccgactatggtgtaagctggattgccagcctccacgaaa gggtctggagtggctgggagtaatatggggtagtgaaacc acatactataattcagctctcaaatccagactgaccatcatc	Axicabtagene ciloleucel CD19 CAR nucleotide sequence

SEQ ID NO:	Sequence	Description
	aaggacaactccaagagccaagtttcttaaaaatgaaca gtctgcaaactgatgacacagccatttactactgtgcaaaa cattattactacggtggtagctatgctatggactactggggtc aaggaacctcagtcaccgtctcctcagcgccgcaattga agttatgtatcctcctctacctagacaatgagaagagca atggaaccattatccatgtgaaagggaaacaccttgtcca agtcccctatttcccgaccttctaagccctttgggtgctggt ggtggttgggggagtcctggctgctatagcttgcagtaac agtggccttattatttctgggtgaggagtaagaggagcag gctcctgcacagtgactacatgaacatgactccccgccgc cccgggcccacccgcaagcattaccagccctatgccccca ccacgcgacttcgcagcctatcgctccagagtgaagttca gcaggagcgagacgccccgcgtaccagcaggggcca gaaccagctctataacgagctcaatctaggacgaagaga ggagtacgatgttttgacaagagacgtggccgggaccct gagatggggggaaagccgagaaggaagaaccctcagg aaggcctgtacaatgaactgcagaaagataagatggcgg aggcctacagtgagattgggatgaaaggcgagcgccgg aggggcaaggggcacgatggccttaccaggggtctcagt acagccaccaaggacacctacgacgcccttcacatgcag gcctgccccctcgc	
1136	MLLLVTSLLLCELPHPAFLLIPDIQMTQTTS SLSASLGDRVTISCRASQDISKYLNWYQQ KPDGTVKLLIYHTSRLHSGVPSRFSGSGS GTDYSLTISNLEQEDIATYFCQQGNTLPYT FGGGTKLEITGSTSGSGKPGSGEGSTKGE VKLQESGPGLVAPSQSLSVTCTVSGVSLP DYGVSWIRQPPRKGLEWLGVWGETTY YNSALKSRLTIKDNSKSQVFLKMNSLQTD DTAIYYCAKHYYYGGSYAMDYWGQGTSTV TVSSAAAIEVMYPPPYLDNEKSNGTIIHVK GKHLCPSPFLPGPSKPFWWLVVGGVLAC YSLLVTVAFIIFWVRSKRSLHSDYMNMT PRRPGPTRKHYQPYAPPRDFAAYRSRVK FSRSADAPAYQQGQNQLYNELNLGRREE YDVLDKRRGRDPEMGGKPRRKNPQEGLY NELQKDKMAEAYSEIGMKGERRRGKGHD GLYQGLSTATKDTYDALHMQALPPR	Axicabtagene ciloleucel CD19 CAR amino acid sequence

Table G. Annotation of tisagenlecleucel CD19 CAR sequences

Feature	Nucleotide Sequence Position	Amino Acid Sequence Position
CD8 α signal peptide	1-63	1-21
FMC63 scFv	64-789	22-263

Feature	Nucleotide Sequence Position	Amino Acid Sequence Position
(V _L -3xG ₄ S linker-V _H)		
CD8 α hinge domain	790-924	264-308
CD8 α transmembrane domain	925-996	309-332
4-1BB costimulatory domain	997-1122	333-374
CD3 ζ signaling domain	1123-1458	375-486

Table H. Annotation of lisocabtagene maraleucel CD19 CAR sequences

Feature	Nucleotide Sequence Position	Amino Acid Sequence Position
GMCSFR- α signal peptide	1-66	1-22
FMC63 scFv (V _L -Whitlow linker-V _H)	67-801	23-267
IgG4 hinge domain	802-837	268-279
CD28 transmembrane domain	838-921	280-307
4-1BB costimulatory domain	922-1047	308-349
CD3 ζ signaling domain	1048-1383	350-461

Table I. Annotation of axicabtagene ciloleucel CD19 CAR sequences

Feature	Nucleotide Sequence Position	Amino Acid Sequence Position
CSF2RA signal peptide	1-66	1-22
FMC63 scFv (V _L -Whitlow linker-V _H)	67-801	23-267
CD28 hinge domain	802-927	268-309
CD28 transmembrane domain	928-1008	310-336
CD28 costimulatory domain	1009-1131	337-377
CD3 ζ signaling domain	1132-1467	378-489

In some embodiments, the polycistronic vector comprises an expression cassette that contains a nucleotide sequence encoding CD19 CAR as set forth in SEQ ID NO: 1131, 1133, or 1135, or at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the nucleotide sequence set forth in SEQ ID NO: 1131, 1133, or

1135. The encoded CD19 CAR has a corresponding amino acid sequence set forth in SEQ ID NO: 1132, 1134, or 1136, respectively, is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in of SEQ ID NO: 1132, 1134, or 1136, respectively.

CD20 CAR

In some embodiments, the CAR is a CD20 CAR ("CD20-CAR"), and in these embodiments, the polycistronic vector comprises an expression cassette that contains a nucleotide sequence encoding a CD20 CAR. CD20 is an antigen found on the surface of B cells as early at the pro-B phase and progressively at increasing levels until B cell maturity, as well as on the cells of most B-cell neoplasms. CD20 positive cells are also sometimes found in cases of Hodgkins disease, myeloma, and thymoma. In some embodiments, the CD20 CAR may comprise a signal peptide, an extracellular binding domain that specifically binds CD20, a hinge domain, a transmembrane domain, an intracellular costimulatory domain, and/or an intracellular signaling domain in tandem.

In some embodiments, the signal peptide of the CD20 CAR comprises a CD8 α signal peptide. In some embodiments, the CD8 α signal peptide comprises or consists of an amino acid sequence set forth in SEQ ID NO:1106 or an amino acid sequence that is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in of SEQ ID NO:1106. In some embodiments, the signal peptide comprises an IgK signal peptide. In some embodiments, the IgK signal peptide comprises or consists of an amino acid sequence set forth in SEQ ID NO:1107 or an amino acid sequence that is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in of SEQ ID NO:1107. In some embodiments, the signal peptide comprises a GMCSFR- α or CSF2RA signal peptide. In some embodiments, the GMCSFR- α or CSF2RA signal peptide comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1108 or an amino acid sequence that is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least

96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in of SEQ ID NO: 1108.

In some embodiments, the extracellular binding domain of the CD20 CAR is specific to CD20, for example, human CD20. The extracellular binding domain of the CD20 CAR can be codon-optimized for expression in a host cell or to have variant sequences to increase functions of the extracellular binding domain. In some embodiments, the extracellular binding domain comprises an immunogenically active portion of an immunoglobulin molecule, for example, an scFv.

In some embodiments, the extracellular binding domain of the CD20 CAR is derived from an antibody specific to CD20, including, for example, Leu16, IF5, 1.5.3, rituximab, obinutuzumab, ibritumomab, ofatumumab, tositumumab, odronextamab, veltuzumab, ublituximab, and ocrelizumab. In any of these embodiments, the extracellular binding domain of the CD20 CAR can comprise or consist of the V_H, the V_L, and/or one or more CDRs of any of the antibodies.

In some embodiments, the extracellular binding domain of the CD20 CAR comprises an scFv derived from the Leu16 monoclonal antibody, which comprises the heavy chain variable region (V_H) and the light chain variable region (V_L) of Leu16 connected by a linker. See Wu et al., Protein Engineering. 14(12):1025-1033 (2001). In some embodiments, the linker is a 3xG4S linker (SEQ ID NO: 1130). In other embodiments, the linker is a Whitlow linker as described herein. In some embodiments, the amino acid sequences of different portions of the entire Leu16-derived scFv (also referred to as Leu16 scFv) and its different portions are provided in **Table J** below. In some embodiments, the CD20-specific scFv comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1137, 1138, or 1142, or an amino acid sequence that is at least 80% identical (*e.g.*, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in of SEQ ID NO: 1137, 1138, or 1142. In some embodiments, the CD20-specific scFv may comprise one or more CDRs having amino acid sequences set forth in SEQ ID NOs: 1139-1141, 1143 and 1144. In some embodiments, the CD20-specific scFv may comprise a light chain with one or more CDRs having amino acid sequences set forth in SEQ ID NOs: 1139-1141. In some embodiments, the CD20-specific scFv may comprise a heavy chain with one or more CDRs having amino acid sequences set forth in SEQ

ID NOs: 1143-1144. In any of these embodiments, the CD20-specific scFv may comprise one or more CDRs comprising one or more amino acid substitutions, or comprising a sequence that is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical), to any of the sequences identified. In some embodiments, the extracellular binding domain of the CD20 CAR comprises or consists of the one or more CDRs as described herein.

Table J. Exemplary sequences of anti-CD20 scFv and components

SEQ ID NO:	Amino Acid Sequence	Description
1137	DIVLTQSPAILSASPGEKVTMTTCRA SSSVNYMDWYQKKPGSSPKPWIYA TSNLASGVPARFSGSGSGTSYSLTI SRVEAEDAATYYCQQWSFNPPTFG GGTKLEIKGSTSGSGKPGSGEGST KGEVQLQQSGAELVKPGASVKMSC KASGYTFTSYNMHWWKQTPGQGL EWIGAIYPGNGDTSYNQKFKGKATL TADKSSSTAYMQLSSLTSEDSADY YCARSNYYGSSYWFFDVWGAGTT VTVSS	Anti-CD20 Leu16 scFv entire sequence, with Whitlow linker
1138	DIVLTQSPAILSASPGEKVTMTTCRA SSSVNYMDWYQKKPGSSPKPWIYA TSNLASGVPARFSGSGSGTSYSLTI SRVEAEDAATYYCQQWSFNPPTFG GGTKLEIK	Anti-CD20 Leu16 scFv light chain variable region
1139	RASSSVNYMD	Anti-CD20 Leu16 scFv light chain CDR1
1140	ATSNLAS	Anti-CD20 Leu16 scFv light chain CDR2
1141	QQWSFNPPT	Anti-CD20 Leu16 scFv light chain CDR3
1142	EVQLQQSGAELVKPGASVKMSCKA SGYTFTSYNMHWWKQTPGQGLEWI GAIYPGNGDTSYNQKFKGKATLTA DKSSSTAYMQLSSLTSEDSADYYC ARSNYYGSSYWFFDVWGAGTTVT VSS	Anti-CD20 Leu16 scFv heavy chain
1143	SYNMH	Anti-CD20 Leu16 scFv heavy chain CDR1

SEQ ID NO:	Amino Acid Sequence	Description
1144	AIYPGNGDTSYNQKFKG	Anti-CD20 Leu16 scFv heavy chain CDR2

In some embodiments, the hinge domain of the CD20 CAR comprises a CD8 α hinge domain, for example, a human CD8 α hinge domain. In some embodiments, the CD8 α hinge domain comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1109 or an amino acid sequence that is at least 80% identical (*e.g.*, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in of SEQ ID NO: 1109. In some embodiments, the hinge domain comprises a CD28 hinge domain, for example, a human CD28 hinge domain. In some embodiments, the CD28 hinge domain comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1110 or an amino acid sequence that is at least 80% identical (*e.g.*, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in of SEQ ID NO: 1110. In some embodiments, the hinge domain comprises an IgG4 hinge domain, for example, a human IgG4 hinge domain. In some embodiments, the IgG4 hinge domain comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1111 or SEQ ID NO: 1112, or an amino acid sequence that is at least 80% identical (*e.g.*, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in of SEQ ID NO: 1111 or SEQ ID NO: 1112. In some embodiments, the hinge domain comprises a IgG4 hinge-Ch2-Ch3 domain, for example, a human IgG4 hinge-Ch2-Ch3 domain. In some embodiments, the IgG4 hinge-Ch2-Ch3 domain comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1113 or an amino acid sequence that is at least 80% identical (*e.g.*, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in of SEQ ID NO: 1113.

In some embodiments, the transmembrane domain of the CD20 CAR comprises a CD8 α transmembrane domain, for example, a human CD8 α transmembrane domain. In some embodiments, the CD8 α transmembrane domain comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1114 or an amino acid

sequence that is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in SEQ ID NO: 1114. In some embodiments, the transmembrane domain comprises a CD28 transmembrane domain, for example, a human CD28 transmembrane domain. In some embodiments, the CD28 transmembrane domain comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1115 or an amino acid sequence that is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in SEQ ID NO: 1115.

In some embodiments, the intracellular costimulatory domain of the CD20 CAR comprises a 4-1BB costimulatory domain, for example, a human 4-1BB costimulatory domain. In some embodiments, the 4-1BB costimulatory domain comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1116 or an amino acid sequence that is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in SEQ ID NO: 1116. In some embodiments, the intracellular costimulatory domain comprises a CD28 costimulatory domain, for example, a human CD28 costimulatory domain. In some embodiments, the CD28 costimulatory domain comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1117 or an amino acid sequence that is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in SEQ ID NO: 1117.

In some embodiments, the intracellular signaling domain of the CD20 CAR comprises a CD3 zeta (ζ) signaling domain, for example, a human CD3 ζ signaling domain. In some embodiments, the CD3 ζ signaling domain comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1118 or an amino acid sequence that is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in SEQ ID NO: 1118.

In some embodiments, the polycistronic vector comprises an expression cassette that contains a nucleotide sequence encoding a CD20 CAR, including, for example,

a CD20 CAR comprising the CD20-specific scFv having sequences set forth in SEQ ID NO: 1137, the CD8 α hinge domain of SEQ ID NO: 1109, the CD8 α transmembrane domain of SEQ ID NO: 1114, the 4-1BB costimulatory domain of SEQ ID NO: 1116, the CD3 ζ signaling domain of SEQ ID NO: 1118, and/or variants (i.e., having a sequence that is at least 80% identical, for example, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99 identical to the disclosed sequence) thereof.

In some embodiments, the polycistronic vector comprises an expression cassette that contains a nucleotide sequence encoding a CD20 CAR, including, for example, a CD20 CAR comprising the CD20-specific scFv having sequences set forth in SEQ ID NO: 1137, the CD28 hinge domain of SEQ ID NO: 1110, the CD8 α transmembrane domain of SEQ ID NO: 1114, the 4-1BB costimulatory domain of SEQ ID NO: 1116, the CD3 ζ signaling domain of SEQ ID NO: 1118, and/or variants (i.e., having a sequence that is at least 80% identical, for example, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99 identical to the disclosed sequence) thereof.

In some embodiments, the polycistronic vector comprises an expression cassette that contains a nucleotide sequence encoding a CD20 CAR, including, for example, a CD20 CAR comprising the CD20-specific scFv having sequences set forth in SEQ ID NO: 1137, the IgG4 hinge domain of SEQ ID NO: 1111 or SEQ ID NO: 1112, the CD8 α transmembrane domain of SEQ ID NO: 1114, the 4-1BB costimulatory domain of SEQ ID NO: 1116, the CD3 ζ signaling domain of SEQ ID NO: 1118, and/or variants (i.e., having a sequence that is at least 80% identical, for example, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99 identical to the disclosed sequence) thereof.

In some embodiments, the polycistronic vector comprises an expression cassette that contains a nucleotide sequence encoding a CD20 CAR, including, for example, a CD20 CAR comprising the CD20-specific scFv having sequences set forth in SEQ ID NO: 1137, the CD8 α hinge domain of SEQ ID NO: 1109, the CD28 transmembrane domain of SEQ ID NO: 1115, the 4-1BB costimulatory domain of SEQ ID NO: 1116, the CD3 ζ signaling domain of SEQ ID NO: 1118, and/or variants (i.e., having a sequence that is at least 80% identical, for example, at least 80%, at

least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99 identical to the disclosed sequence) thereof.

In some embodiments, the polycistronic vector comprises an expression cassette that contains a nucleotide sequence encoding a CD20 CAR, including, for example, a CD20 CAR comprising the CD20-specific scFv having sequences set forth in SEQ ID NO: 1137, the CD28 hinge domain of SEQ ID NO: 1110, the CD28 transmembrane domain of SEQ ID NO: 1115, the 4-1BB costimulatory domain of SEQ ID NO: 1116, the CD3 ζ signaling domain of SEQ ID NO: 1118, and/or variants (i.e., having a sequence that is at least 80% identical, for example, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99 identical to the disclosed sequence) thereof.

In some embodiments, the polycistronic vector comprises an expression cassette that contains a nucleotide sequence encoding a CD20 CAR, including, for example, a CD20 CAR comprising the CD20-specific scFv having sequences set forth in SEQ ID NO: 1137, the IgG4 hinge domain of SEQ ID NO: 1111 or SEQ ID NO: 1112, the CD28 transmembrane domain of SEQ ID NO: 1115, the 4-1BB costimulatory domain of SEQ ID NO: 1116, the CD3 ζ signaling domain of SEQ ID NO: 1118, and/or variants (i.e., having a sequence that is at least 80% identical, for example, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99 identical to the disclosed sequence) thereof.

CD22 CAR

In some embodiments, the CAR is a CD22 CAR ("CD22-CAR"), and in these embodiments, the polycistronic vector comprises an expression cassette that contains a nucleotide sequence encoding a CD22 CAR. CD22, which is a transmembrane protein found mostly on the surface of mature B cells that functions as an inhibitory receptor for B cell receptor (BCR) signaling. CD22 is expressed in 60-70% of B cell lymphomas and leukemias (e.g., B-chronic lymphocytic leukemia, hairy cell leukemia, acute lymphocytic leukemia (ALL), and Burkitt's lymphoma) and is not present on the cell surface in early stages of B cell development or on stem cells. In some embodiments, the CD22 CAR may comprise a signal peptide, an extracellular binding domain that specifically binds CD22, a hinge domain, a transmembrane domain, an intracellular costimulatory domain, and/or an intracellular signaling domain in tandem.

In some embodiments, the signal peptide of the CD22 CAR comprises a CD8 α signal peptide. In some embodiments, the CD8 α signal peptide comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1106 or an amino acid sequence that is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in of SEQ ID NO: 1106. In some embodiments, the signal peptide comprises an IgK signal peptide. In some embodiments, the IgK signal peptide comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1107 or an amino acid sequence that is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in of SEQ ID NO: 1107. In some embodiments, the signal peptide comprises a GMCSFR- α or CSF2RA signal peptide. In some embodiments, the GMCSFR- α or CSF2RA signal peptide comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1108 or an amino acid sequence that is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in of SEQ ID NO: 1108.

In some embodiments, the extracellular binding domain of the CD22 CAR is specific to CD22, for example, human CD22. The extracellular binding domain of the CD22 CAR can be codon-optimized for expression in a host cell or to have variant sequences to increase functions of the extracellular binding domain. In some embodiments, the extracellular binding domain comprises an immunogenically active portion of an immunoglobulin molecule, for example, an scFv.

In some embodiments, the extracellular binding domain of the CD22 CAR is derived from an antibody specific to CD22, including, for example, SM03, inotuzumab, epratuzumab, moxetumomab, and pinatuzumab. In any of these embodiments, the extracellular binding domain of the CD22 CAR can comprise or consist of the V_H, the V_L, and/or one or more CDRs of any of the antibodies.

In some embodiments, the extracellular binding domain of the CD22 CAR comprises an scFv derived from the m971 monoclonal antibody (m971), which comprises the heavy chain variable region (V_H) and the light chain variable region (V_L) of m971 connected by a linker. In some embodiments, the linker is a 3xG4S linker. In other

embodiments, the Whitlow linker may be used instead. In some embodiments, the amino acid sequences of the entire m971-derived scFv (also referred to as m971 scFv) and its different portions are provided in **Table K** below. In some embodiments, the CD22-specific scFv comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1145, 1146, or 1150, or an amino acid sequence that is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in of SEQ ID NO: 1145, 1146, or 1150. In some embodiments, the CD22-specific scFv may comprise one or more CDRs having amino acid sequences set forth in SEQ ID NOs: 1147-1149 and 1151-1153. In some embodiments, the CD22-specific scFv may comprise a heavy chain with one or more CDRs having amino acid sequences set forth in SEQ ID NOs: 1147-1149. In some embodiments, the CD22-specific scFv may comprise a light chain with one or more CDRs having amino acid sequences set forth in SEQ ID NOs: 1151-1153. In any of these embodiments, the CD22-specific scFv may comprise one or more CDRs comprising one or more amino acid substitutions, or comprising a sequence that is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical), to any of the sequences identified. In some embodiments, the extracellular binding domain of the CD22 CAR comprises or consists of the one or more CDRs as described herein. In some embodiments, the extracellular binding domain of the CD22 CAR comprises an scFv derived from m971-L7, which is an affinity matured variant of m971 with significantly improved CD22 binding affinity compared to the parental antibody m971 (improved from about 2 nM to less than 50 pM). In some embodiments, the scFv derived from m971-L7 comprises the V_H and the V_L of m971-L7 connected by a 3xG4S linker (SEQ ID NO: 1130). In other embodiments, the Whitlow linker may be used instead. In some embodiments, the amino acid sequences of the entire m971-L7-derived scFv (also referred to as m971-L7 scFv) and its different portions are provided in **Table K** below. In some embodiments, the CD22-specific scFv comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1154, 1155, or 1159, or an amino acid sequence that is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in of SEQ

ID NO: 1154, 1155, or 1159. In some embodiments, the CD22-specific scFv may comprise one or more CDRs having amino acid sequences set forth in SEQ ID NOs: 1156-1158 and 1160-1162. In some embodiments, the CD22-specific scFv may comprise a heavy chain with one or more CDRs having amino acid sequences set forth in SEQ ID NOs: 1156-1158. In some embodiments, the CD22-specific scFv may comprise a light chain with one or more CDRs having amino acid sequences set forth in SEQ ID NOs: 1160-1162. In any of these embodiments, the CD22-specific scFv may comprise one or more CDRs comprising one or more amino acid substitutions, or comprising a sequence that is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical), to any of the sequences identified. In some embodiments, the extracellular binding domain of the CD22 CAR comprises or consists of the one or more CDRs as described herein.

Table K. Exemplary sequences of anti-CD22 scFv and components

SEQ ID NO:	Amino Acid Sequence	Description
1145	QVQLQQSGPGLVKPSQTLSTCAIS GDSVSSNSAAWNWIRQSPSRGLE WLGRTYYRSKWYNDYAVSVKSRITI NPDTSKNQFSLQLNSVTPEDTAVY YCAREVTGDLEDAFDIWGQGTMT VSSGGGGSGGGGSGGGGSDIQMT QSPSSLASVGDRTITCRASQTIW SYLNWYQQRPGKAPNLLIYAASSL QSGVPSRFSGRSGTDFLTITSSSLQ AEDFATYYCQQSYSIPQTFGQGTKL EIK	Anti-CD22 m971 scFv entire sequence, with 3xG4S linker (SEQ ID NO: 1130)
1146	QVQLQQSGPGLVKPSQTLSTCAIS GDSVSSNSAAWNWIRQSPSRGLE WLGRTYYRSKWYNDYAVSVKSRITI NPDTSKNQFSLQLNSVTPEDTAVY YCAREVTGDLEDAFDIWGQGTMT VSS	Anti-CD22 m971 scFv heavy chain variable region
1147	GDSVSSNSAA	Anti-CD22 m971 scFv heavy chain CDR1
1148	YYRSKWYN	Anti-CD22 m971 scFv heavy chain CDR2
1149	AREVTGDLEDAFDI	Anti-CD22 m971 scFv heavy chain CDR3

SEQ ID NO:	Amino Acid Sequence	Description
1150	DIQMTQSPSSLSASVGDRVTITCRA SQTIEWSYLNWYQQRPGKAPNLLIYA ASSLQSGVPSRFSGRSGTDFTLTI SSLQAEDFATYYCQQSYSIPQTFG QGKLEIK	Anti-CD22 m971 scFv light chain
1151	QTIWSY	Anti-CD22 m971 scFv light chain CDR1
1152	AAS	Anti-CD22 m971 scFv light chain CDR2
1153	QQSYSIPQT	Anti-CD22 m971 scFv light chain CDR3
1154	QVQLQQSGPGMVKPSQTLSTCAI SGDSVSSNSVAWNWIRQSPSRGLE WLGRTYYRSTWYNDYAVSMKSRIT INPDNKNQFSLQLNSVTPEDTAVY YCAREVTGDLEDAFDIWGQGMVT VSSGGGGSGGGGSGGGGSDIQMI QSPSSLSASVGDRVTITCRASQTIW SYLNWYRQRPGEAPNLLIYAASSLQ SGVPSRFSGRSGTDFTLTISSLQA EDFATYYCQQSYSIPQTFGQGKLE IK	Anti-CD22 m971-L7 scFv entire sequence, with 3xG4S linker (SEQ ID NO: 1130)
1155	QVQLQQSGPGMVKPSQTLSTCAI SGDSVSSNSVAWNWIRQSPSRGLE WLGRTYYRSTWYNDYAVSMKSRIT INPDNKNQFSLQLNSVTPEDTAVY YCAREVTGDLEDAFDIWGQGMVT VSS	Anti-CD22 m971-L7 scFv heavy chain variable region
1156	GDSVSSNSVA	Anti-CD22 m971-L7 scFv heavy chain CDR1
1157	TYRSTWYN	Anti-CD22 m971-L7 scFv heavy chain CDR2
1158	AREVTGDLEDAFDI	Anti-CD22 m971-L7 scFv heavy chain CDR3
1159	DIQMIQSPSSLSASVGDRVTITCRA SQTIEWSYLNWYRQRPGEAPNLLIYA ASSLQSGVPSRFSGRSGTDFTLTI SSLQAEDFATYYCQQSYSIPQTFG QGKLEIK	Anti-CD22 m971-L7 scFv light chain variable region

SEQ ID NO:	Amino Acid Sequence	Description
1160	QTIWSY	Anti-CD22 m971-L7 scFv light chain CDR1
1161	AAS	Anti-CD22 m971-L7 scFv light chain CDR2
1162	QQSYSIPQT	Anti-CD22 m971-L7 scFv light chain CDR3

In some embodiments, the extracellular binding domain of the CD22 CAR comprises immunotoxins HA22 or BL22. Immunotoxins BL22 and HA22 are therapeutic agents that comprise an scFv specific for CD22 fused to a bacterial toxin, and thus can bind to the surface of the cancer cells that express CD22 and kill the cancer cells. BL22 comprises a dsFv of an anti-CD22 antibody, RFB4, fused to a 38-kDa truncated form of *Pseudomonas* exotoxin A (Bang et al., Clin. Cancer Res., 11:1545-50 (2005)). HA22 (CAT8015, moxetumomab pasudotox) is a mutated, higher affinity version of BL22 (Ho et al., J. Biol. Chem., 280(1): 607-17 (2005)). Suitable sequences of antigen binding domains of HA22 and BL22 specific to CD22 are disclosed in, for example, U.S. Patent Nos. 7,541,034; 7,355,012; and 7,982,011, which are hereby incorporated by reference in their entirety.

In some embodiments, the hinge domain of the CD22 CAR comprises a CD8 α hinge domain, for example, a human CD8 α hinge domain. In some embodiments, the CD8 α hinge domain comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1109 or an amino acid sequence that is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in of SEQ ID NO: 1109. In some embodiments, the hinge domain comprises a CD28 hinge domain, for example, a human CD28 hinge domain. In some embodiments, the CD28 hinge domain comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1110 or an amino acid sequence that is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in of SEQ ID NO: 1110. In some embodiments, the hinge domain comprises an IgG4 hinge domain, for example, a human IgG4 hinge domain. In some embodiments, the IgG4 hinge domain comprises or consists of an amino acid sequence set forth in

SEQ ID NO: 1111 or SEQ ID NO: 1112, or an amino acid sequence that is at least 80% identical (*e.g.*, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in of SEQ ID NO: 1111 or SEQ ID NO: 1112. In some embodiments, the hinge domain comprises a IgG4 hinge-Ch2-Ch3 domain, for example, a human IgG4 hinge-Ch2-Ch3 domain. In some embodiments, the IgG4 hinge-Ch2-Ch3 domain comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1113 or an amino acid sequence that is at least 80% identical (*e.g.*, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in of SEQ ID NO: 1113.

In some embodiments, the transmembrane domain of the CD22 CAR comprises a CD8 α transmembrane domain, for example, a human CD8 α transmembrane domain. In some embodiments, the CD8 α transmembrane domain comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1114 or an amino acid sequence that is at least 80% identical (*e.g.*, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in SEQ ID NO: 1114. In some embodiments, the transmembrane domain comprises a CD28 transmembrane domain, for example, a human CD28 transmembrane domain. In some embodiments, the CD28 transmembrane domain comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1115 or an amino acid sequence that is at least 80% identical (*e.g.*, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in SEQ ID NO: 1115.

In some embodiments, the intracellular costimulatory domain of the CD22 CAR comprises a 4-1BB costimulatory domain, for example, a human 4-1BB costimulatory domain. In some embodiments, the 4-1BB costimulatory domain comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1116 or an amino acid sequence that is at least 80% identical (*e.g.*, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in SEQ ID NO: 1116. In some embodiments, the intracellular costimulatory domain comprises a CD28

costimulatory domain, for example, a human CD28 costimulatory domain. In some embodiments, the CD28 costimulatory domain comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1117 or an amino acid sequence that is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in SEQ ID NO: 1117.

In some embodiments, the intracellular signaling domain of the CD22 CAR comprises a CD3 zeta (ζ) signaling domain, for example, a human CD3 ζ signaling domain. In some embodiments, the CD3 ζ signaling domain comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1118 or an amino acid sequence that is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in SEQ ID NO: 1118.

In some embodiments, the polycistronic vector comprises an expression cassette that contains a nucleotide sequence encoding a CD22 CAR, including, for example, a CD22 CAR comprising the CD22-specific scFv having sequences set forth in SEQ ID NO: 1145 or SEQ ID NO: 1154, the CD8 α hinge domain of SEQ ID NO: 9, the CD8 α transmembrane domain of SEQ ID NO: 1114, the 4-1BB costimulatory domain of SEQ ID NO: 1116, the CD3 ζ signaling domain of SEQ ID NO: 1118, and/or variants (i.e., having a sequence that is at least 80% identical, for example, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99 identical to the disclosed sequence) thereof.

In some embodiments, the polycistronic vector comprises an expression cassette that contains a nucleotide sequence encoding a CD22 CAR, including, for example, a CD22 CAR comprising the CD22-specific scFv having sequences set forth in SEQ ID NO: 1145 or SEQ ID NO: 1154, the CD28 hinge domain of SEQ ID NO: 1110, the CD8 α transmembrane domain of SEQ ID NO: 1114, the 4-1BB costimulatory domain of SEQ ID NO: 1116, the CD3 ζ signaling domain of SEQ ID NO: 1118, and/or variants (i.e., having a sequence that is at least 80% identical, for example, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99 identical to the disclosed sequence) thereof.

In some embodiments, the polycistronic vector comprises an expression cassette that contains a nucleotide sequence encoding a CD22 CAR, including, for example,

a CD22 CAR comprising the CD22-specific scFv having sequences set forth in SEQ ID NO: 1145 or SEQ ID NO: 1154, the IgG4 hinge domain of SEQ ID NO: 1111 or SEQ ID NO: 1112, the CD8 α transmembrane domain of SEQ ID NO: 1114, the 4-1BB costimulatory domain of SEQ ID NO: 1116, the CD3 ζ signaling domain of SEQ ID NO: 1118, and/or variants (i.e., having a sequence that is at least 80% identical, for example, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99 identical to the disclosed sequence) thereof. In some embodiments, the polycistronic vector comprises an expression cassette that contains a nucleotide sequence encoding a CD22 CAR, including, for example, a CD22 CAR comprising the CD22-specific scFv having sequences set forth in SEQ ID NO: 1145 or SEQ ID NO: 1154, the CD8 α hinge domain of SEQ ID NO: 1114, the CD28 transmembrane domain of SEQ ID NO: 1115, the 4-1BB costimulatory domain of SEQ ID NO: 1116, the CD3 ζ signaling domain of SEQ ID NO: 1118, and/or variants (i.e., having a sequence that is at least 80% identical, for example, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99 identical to the disclosed sequence) thereof. In some embodiments, the polycistronic vector comprises an expression cassette that contains a nucleotide sequence encoding a CD22 CAR, including, for example, a CD22 CAR comprising the CD22-specific scFv having sequences set forth in SEQ ID NO: 1145 or SEQ ID NO: 1154, the CD28 hinge domain of SEQ ID NO: 1110, the CD28 transmembrane domain of SEQ ID NO: 1115, the 4-1BB costimulatory domain of SEQ ID NO: 1116, the CD3 ζ signaling domain of SEQ ID NO: 1118, and/or variants (i.e., having a sequence that is at least 80% identical, for example, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99 identical to the disclosed sequence) thereof. In some embodiments, the polycistronic vector comprises an expression cassette that contains a nucleotide sequence encoding a CD22 CAR, including, for example, a CD22 CAR comprising the CD22-specific scFv having sequences set forth in SEQ ID NO: 1145 or SEQ ID NO: 1154, the IgG4 hinge domain of SEQ ID NO: 1111 or SEQ ID NO: 1112, the CD28 transmembrane domain of SEQ ID NO: 1115, the 4-1BB costimulatory domain of SEQ ID NO: 1116, the CD3 ζ signaling domain of SEQ ID NO: 1118, and/or variants (i.e., having a sequence that is at least 80% identical,

for example, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99 identical to the disclosed sequence) thereof.

BCMA CAR

In some embodiments, the CAR is a BCMA CAR ("BCMA-CAR"), and in these embodiments, the polycistronic vector comprises an expression cassette that contains a nucleotide sequence encoding a BCMA CAR. BCMA is a tumor necrosis family receptor (TNFR) member expressed on cells of the B cell lineage, with the highest expression on terminally differentiated B cells or mature B lymphocytes. BCMA is involved in mediating the survival of plasma cells for maintaining long-term humoral immunity. The expression of BCMA has been recently linked to a number of cancers, such as multiple myeloma, Hodgkin's and non-Hodgkin's lymphoma, various leukemias, and glioblastoma. In some embodiments, the BCMA CAR may comprise a signal peptide, an extracellular binding domain that specifically binds BCMA, a hinge domain, a transmembrane domain, an intracellular costimulatory domain, and/or an intracellular signaling domain in tandem.

In some embodiments, the signal peptide of the BCMA CAR comprises a CD8 α signal peptide. In some embodiments, the CD8 α signal peptide comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1106 or an amino acid sequence that is at least 80% identical (*e.g.*, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in of SEQ ID NO: 1106. In some embodiments, the signal peptide comprises an IgK signal peptide. In some embodiments, the IgK signal peptide comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1107 or an amino acid sequence that is at least 80% identical (*e.g.*, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in of SEQ ID NO: 1107. In some embodiments, the signal peptide comprises a GMCSFR- α or CSF2RA signal peptide. In some embodiments, the GMCSFR- α or CSF2RA signal peptide comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1108 or an amino acid sequence that is at least 80% identical (*e.g.*, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in of SEQ ID NO: 1108.

In some embodiments, the extracellular binding domain of the BCMA CAR is specific to BCMA, for example, human BCMA. The extracellular binding domain of the BCMA CAR can be codon-optimized for expression in a host cell or to have variant sequences to increase functions of the extracellular binding domain.

In some embodiments, the extracellular binding domain comprises an immunogenically active portion of an immunoglobulin molecule, for example, an scFv. In some embodiments, the extracellular binding domain of the BCMA CAR is derived from an antibody specific to BCMA, including, for example, belantamab, erlanatamab, teclistamab, LCAR-B38M, and ciltacabtagene. In any of these embodiments, the extracellular binding domain of the BCMA CAR can comprise or consist of the V_H, the V_L, and/or one or more CDRs of any of the antibodies.

In some embodiments, the extracellular binding domain of the BCMA CAR comprises an scFv derived from C11D5.3, a murine monoclonal antibody as described in Carpenter et al., Clin. Cancer Res. 19(8):2048-2060 (2013). See also PCT Application Publication No. WO2010/104949. The C11D5.3-derived scFv may comprise the heavy chain variable region (V_H) and the light chain variable region (V_L) of C11D5.3 connected by the Whitlow linker, the amino acid sequences of which is provided in **Table L** below. In some embodiments, the BCMA-specific extracellular binding domain comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1163, 1164, or 1168, or an amino acid sequence that is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in of SEQ ID NO: 1163, 1164, or 1168. In some embodiments, the BCMA-specific extracellular binding domain may comprise one or more CDRs having amino acid sequences set forth in SEQ ID NOs: 1165-1167 and 1169-1171. In some embodiments, the BCMA-specific extracellular binding domain may comprise a light chain with one or more CDRs having amino acid sequences set forth in SEQ ID NOs: 1165-1167. In some embodiments, the BCMA-specific extracellular binding domain may comprise a heavy chain with one or more CDRs having amino acid sequences set forth in SEQ ID NOs: 1169-1171. In any of these embodiments, the BCMA-specific scFv may comprise one or more CDRs comprising one or more amino acid substitutions, or comprising a sequence that is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least

97%, at least 98%, at least 99%, or 100% identical), to any of the sequences identified. In some embodiments, the extracellular binding domain of the BCMA CAR comprises or consists of the one or more CDRs as described herein.

In some embodiments, the extracellular binding domain of the BCMA CAR comprises an scFv derived from another murine monoclonal antibody, C12A3.2, as described in Carpenter et al., Clin. Cancer Res. 19(8):2048-2060 (2013) and PCT Application Publication No. WO2010/104949, the amino acid sequence of which is also provided in **Table L** below. In some embodiments, the BCMA-specific extracellular binding domain comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1172, 1173, or 1177, or an amino acid sequence that is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in of SEQ ID NO: 1172, 1173, or 1177. In some embodiments, the BCMA-specific extracellular binding domain may comprise one or more CDRs having amino acid sequences set forth in SEQ ID NOs: 1174-1176 and 1178-1180. In some embodiments, the BCMA-specific extracellular binding domain may comprise a light chain with one or more CDRs having amino acid sequences set forth in SEQ ID NOs: 1174-1176. In some embodiments, the BCMA-specific extracellular binding domain may comprise a heavy chain with one or more CDRs having amino acid sequences set forth in SEQ ID NOs: 1178-1180. In any of these embodiments, the BCMA-specific scFv may comprise one or more CDRs comprising one or more amino acid substitutions, or comprising a sequence that is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical), to any of the sequences identified. In some embodiments, the extracellular binding domain of the BCMA CAR comprises or consists of the one or more CDRs as described herein.

In some embodiments, the extracellular binding domain of the BCMA CAR comprises a murine monoclonal antibody with high specificity to human BCMA, referred to as BB2121 in Friedman et al., Hum. Gene Ther. 29(5):585-601 (2018)). See also, PCT Application Publication No. WO2012163805.

In some embodiments, the extracellular binding domain of the BCMA CAR comprises single variable fragments of two heavy chains (VHH) that can bind to two epitopes of BCMA as described in Zhao et al., J. Hematol. Oncol. 11(1):141 (2018),

also referred to as LCAR-B38M. See also, PCT Application Publication No. WO2018/028647.

In some embodiments, the extracellular binding domain of the BCMA CAR comprises a fully human heavy-chain variable domain (FHVH) as described in Lam et al., Nat. Commun. 11(1):283 (2020), also referred to as FHVH33. See also, PCT Application Publication No. WO2019/006072. The amino acid sequences of FHVH33 and its CDRs are provided in **Table L** below. In some embodiments, the BCMA-specific extracellular binding domain comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1181 or an amino acid sequence that is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in of SEQ ID NO: 1181. In some embodiments, the BCMA-specific extracellular binding domain may comprise one or more CDRs having amino acid sequences set forth in SEQ ID NOs: 1182-1184. In any of these embodiments, the BCMA-specific extracellular binding domain may comprise one or more CDRs comprising one or more amino acid substitutions, or comprising a sequence that is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical), to any of the sequences identified. In some embodiments, the extracellular binding domain of the BCMA CAR comprises or consists of the one or more CDRs as described herein. In some embodiments, the extracellular binding domain of the BCMA CAR comprises an scFv derived from CT103A (or CAR0085) as described in U.S. Patent No. 11,026,975 B2, the amino acid sequence of which is provided in **Table L** below. In some embodiments, the BCMA-specific extracellular binding domain comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1218, 1219, or 1223, or an amino acid sequence that is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in of SEQ ID NO: 1218, 1219, or 1223. In some embodiments, the BCMA-specific extracellular binding domain may comprise one or more CDRs having amino acid sequences set forth in SEQ ID NOs: 1220-1222 and 1224-1226. In some embodiments, the BCMA-specific extracellular binding domain may comprise a light chain with one or more CDRs having amino acid sequences set forth in SEQ ID NOs: 1220-1222. In some

embodiments, the BCMA-specific extracellular binding domain may comprise a heavy chain with one or more CDRs having amino acid sequences set forth in SEQ ID NOs: 1224-1226. In any of these embodiments, the BCMA-specific scFv may comprise one or more CDRs comprising one or more amino acid substitutions, or comprising a sequence that is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical), to any of the sequences identified. In some embodiments, the extracellular binding domain of the BCMA CAR comprises or consists of the one or more CDRs as described herein.

Additionally, CARs and binders directed to BCMA have been described in U.S. Application Publication Nos. 2020/0246381 A1 and 2020/0339699 A1, the entire contents of each of which are incorporated by reference herein.

Table L. Exemplary sequences of anti-BCMA binder and components

SEQ ID NO:	Amino Acid Sequence	Description
1163	DIVLTQSPASLAMS LGKRATISCRA SESVSVIGAH LIHWYQQKPGQPPKL LIY LASNLETGVPARFSGSGSGTDF TLTIDPVEEDDVAIYSCLQSRIFPRT FGGGTKLEIKGSTSGSGKPGSGEG STKGQIQLVQSGPELKKPGETVKIS CKASGYTFTDYSINWVKRAPGKGL KWMGWINTETREPAYAYDFRGRFA FSLETSA STAYLQINN LKYEDTATYF CALDYSYAMDYWGQGTSTVTVSS	Anti-BCMA C11D5.3 scFv entire sequence, with Whitlow linker
1164	DIVLTQSPASLAMS LGKRATISCRA SESVSVIGAH LIHWYQQKPGQPPKL LIY LASNLETGVPARFSGSGSGTDF TLTIDPVEEDDVAIYSCLQSRIFPRT FGGGTKLEIK	Anti-BCMA C11D5.3 scFv light chain variable region
1165	RASESVSVIGAH LIH	Anti-BCMA C11D5.3 scFv light chain CDR1
1166	LASNLET	Anti-BCMA C11D5.3 scFv light chain CDR2
1167	LQSRIFPRT	Anti-BCMA C11D5.3 scFv light chain CDR3
1168	QIQLVQSGPELKKPGETVKISCKAS GYTFTDYSINWVKRAPGKGLKWMG WINTETREPAYAYDFRGRFAFSLET	Anti-BCMA C11D5.3 scFv heavy chain variable region

SEQ ID NO:	Amino Acid Sequence	Description
	SASTAYLQINNLYEDTATYFCALD YSYAMDYWGQGTSTVTVSS	
1169	DYSIN	Anti-BCMA C11D5.3 scFv heavy chain CDR1
1170	WINTETREPAYAYDFRG	Anti-BCMA C11D5.3 scFv heavy chain CDR2
1171	DYSYAMDY	Anti-BCMA C11D5.3 scFv heavy chain CDR3
1172	DIVLTQSPPSLAMSLGKRATISCRA SESVTILGSHLIYWYQQKPGQPPTL LIQLASNVQGTGVPARFSGSGSRTDF TLTIDPVEEDDVAVYYCLQSRTIPRT FGGGTKLEIKGSTSGSGKPGSGEG STKGQIQLVQSGPELKKPGETVKIS CKASGYTFRHYSMNWWKQAPGKG LKWMGRINTESGVPIYADDFKGRFA FSVETSASTAYLVINNLYKDEDTASYF CSNDYLYSLDFWGQGTALTVSS	Anti-BCMA C12A3.2 scFv entire sequence, with Whitlow linker
1173	DIVLTQSPPSLAMSLGKRATISCRA SESVTILGSHLIYWYQQKPGQPPTL LIQLASNVQGTGVPARFSGSGSRTDF TLTIDPVEEDDVAVYYCLQSRTIPRT FGGGTKLEIK	Anti-BCMA C12A3.2 scFv light chain variable region
1174	RASESVTILGSHLIY	Anti-BCMA C12A3.2 scFv light chain CDR1
1175	LASNVQT	Anti-BCMA C12A3.2 scFv light chain CDR2
1176	LQSRTIPRT	Anti-BCMA C12A3.2 scFv light chain CDR3
1177	QIQLVQSGPELKKPGETVKISCKAS GYTFRHYSMNWWKQAPGKGLKWM GRINTESGVPIYADDFKGRFAFSVE TSASTAYLVINNLYKDEDTASYFCSN DYLYSLDFWGQGTALTVSS	Anti-BCMA C12A3.2 scFv heavy chain variable region
1178	HYSMN	Anti-BCMA C12A3.2 scFv heavy chain CDR1

SEQ ID NO:	Amino Acid Sequence	Description
1179	RINTESGVPIYADDFKG	Anti-BCMA C12A3.2 scFv heavy chain CDR2
1180	DYLYSLDF	Anti-BCMA C12A3.2 scFv heavy chain CDR3
1181	EVQLLESGGGLVQPGGSLRLSCAA SGFTFSSYAMSWRQAPGKGLEW VSSISGSGDYIYYADSVKGRFTISR ISKNTLYLQMNSLRAEDTAVYYCAK EGTGANSSLADYRGQGTLTVSS	Anti-BCMA FHVH33 entire sequence
1182	GFTFSSYA	Anti-BCMA FHVH33 CDR1
1183	ISGSGDYI	Anti-BCMA FHVH33 CDR2
1184	AKEGTGANSSLADY	Anti-BCMA FHVH33 CDR3
1218	DIQMTQSPSSLSASVGDRVTITCRA SQSISSYLNWYQQKPGKAPKLLIYA ASSLQSGVPSRFSGSGSGTDFTLT SSLQPEDFATYYCQQKYDLLTFGG GTKVEIKGSTSGSGKPGSGEGSTK GQLQLQESGPGLVKPSSETLSLTCTV SGSISSSSYWGWIRQPPGKGLE WIGSISYSGSTYYNPSLKSRVTISVD TSKNQFSLKLSSVTAADTAVYYCAR DRGDTILDVWGQGTMTVTVSS	Anti-BCMA CT103A scFv entire sequence, with Whitlow linker
1219	DIQMTQSPSSLSASVGDRVTITCRA SQSISSYLNWYQQKPGKAPKLLIYA ASSLQSGVPSRFSGSGSGTDFTLT SSLQPEDFATYYCQQKYDLLTFGG GTKVEIK	Anti-BCMA CT103A scFv light chain variable region
1220	QSISSY	Anti-BCMA CT103A scFv light chain CDR1
1221	AAS	Anti-BCMA CT103A scFv light chain CDR2
1222	QQKYDLLT	Anti-BCMA CT103A scFv light chain CDR3
1223	QLQLQESGPGLVKPSSETLSLTCTVS GGSISSSSSYWGWIRQPPGKGLEW IGSISYSGSTYYNPSLKSRVTISVDT	Anti-BCMA CT103A scFv heavy chain variable region

SEQ ID NO:	Amino Acid Sequence	Description
	SKNQFSLKLSSVTAADTAVYYCARD RGDTILDVWGQGTMTVSS	
1224	GGSISSSSYY	Anti-BCMA CT103A scFv heavy chain CDR1
1225	ISYSGST	Anti-BCMA CT103A scFv heavy chain CDR2
1226	ARDRGDTILDV	Anti-BCMA CT103A scFv heavy chain CDR3

In some embodiments, the hinge domain of the BCMA CAR comprises a CD8 α hinge domain, for example, a human CD8 α hinge domain. In some embodiments, the CD8 α hinge domain comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1109 or an amino acid sequence that is at least 80% identical (*e.g.*, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in of SEQ ID NO: 1109. In some embodiments, the hinge domain comprises a CD28 hinge domain, for example, a human CD28 hinge domain. In some embodiments, the CD28 hinge domain comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1110 or an amino acid sequence that is at least 80% identical (*e.g.*, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in of SEQ ID NO: 1110. In some embodiments, the hinge domain comprises an IgG4 hinge domain, for example, a human IgG4 hinge domain. In some embodiments, the IgG4 hinge domain comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1111 or SEQ ID NO: 1112, or an amino acid sequence that is at least 80% identical (*e.g.*, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in of SEQ ID NO: 1111 or SEQ ID NO: 1112. In some embodiments, the hinge domain comprises a IgG4 hinge-Ch2-Ch3 domain, for example, a human IgG4 hinge-Ch2-Ch3 domain. In some embodiments, the IgG4 hinge-Ch2-Ch3 domain comprises or consists of an amino acid sequence set forth in

SEQ ID NO: 1113 or an amino acid sequence that is at least 80% identical (*e.g.*, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in of SEQ ID NO: 1113.

In some embodiments, the transmembrane domain of the BCMA CAR comprises a CD8 α transmembrane domain, for example, a human CD8 α transmembrane domain. In some embodiments, the CD8 α transmembrane domain comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1114 or an amino acid sequence that is at least 80% identical (*e.g.*, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in SEQ ID NO: 1114. In some embodiments, the transmembrane domain comprises a CD28 transmembrane domain, for example, a human CD28 transmembrane domain. In some embodiments, the CD28 transmembrane domain comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1115 or an amino acid sequence that is at least 80% identical (*e.g.*, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in SEQ ID NO: 1115.

In some embodiments, the intracellular costimulatory domain of the BCMA CAR comprises a 4-1BB costimulatory domain, for example, a human 4-1BB costimulatory domain. In some embodiments, the 4-1BB costimulatory domain comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1116 or an amino acid sequence that is at least 80% identical (*e.g.*, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in SEQ ID NO: 1116. In some embodiments, the intracellular costimulatory domain comprises a CD28 costimulatory domain, for example, a human CD28 costimulatory domain. In some embodiments, the CD28 costimulatory domain comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1117 or an amino acid sequence that is at least 80% identical (*e.g.*, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in SEQ ID NO: 1117.

In some embodiments, the intracellular signaling domain of the BCMA CAR comprises a CD3 zeta (ζ) signaling domain, for example, a human CD3 ζ signaling domain. In some embodiments, the CD3 ζ signaling domain comprises or consists of an amino acid sequence set forth in SEQ ID NO: 1118 or an amino acid sequence that is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in SEQ ID NO: 1118.

In some embodiments, the polycistronic vector comprises an expression cassette that contains a nucleotide sequence encoding a BCMA CAR, including, for example, a BCMA CAR comprising any of the BCMA-specific extracellular binding domains as described, the CD8 α hinge domain of SEQ ID NO: 1109, the CD8 α transmembrane domain of SEQ ID NO: 1114, the 4-1BB costimulatory domain of SEQ ID NO: 1116, the CD3 ζ signaling domain of SEQ ID NO: 1118, and/or variants (i.e., having a sequence that is at least 80% identical, for example, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to the disclosed sequence) thereof. In any of these embodiments, the BCMA CAR may additionally comprise a signal peptide (e.g., a CD8 α signal peptide) as described.

In some embodiments, the polycistronic vector comprises an expression cassette that contains a nucleotide sequence encoding a BCMA CAR, including, for example, a BCMA CAR comprising any of the BCMA-specific extracellular binding domains as described, the CD8 α hinge domain of SEQ ID NO: 1109, the CD8 α transmembrane domain of SEQ ID NO: 1114, the CD28 costimulatory domain of SEQ ID NO: 1117, the CD3 ζ signaling domain of SEQ ID NO: 1118, and/or variants (i.e., having a sequence that is at least 80% identical, for example, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to the disclosed sequence) thereof. In any of these embodiments, the BCMA CAR may additionally comprise a signal peptide as described.

In some embodiments, the polycistronic vector comprises an expression cassette that contains a nucleotide sequence encoding a BCMA CAR as set forth in SEQ ID NO: 1227 or is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the nucleotide sequence set forth in SEQ ID NO: 1227 (see **Table M**).

The encoded BCMA CAR has a corresponding amino acid sequence set forth in SEQ ID NO: 1228 or is at least 80% identical (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical) to the amino acid sequence set forth in of SEQ ID NO: 1228, with the following components: CD8 α signal peptide, CT103A scFv (V_L-Whitlow linker-V_H), CD8 α hinge domain, CD8 α transmembrane domain, 4-1BB costimulatory domain, and CD3 ζ signaling domain.

In some embodiments, the polycistronic vector comprises an expression cassette that contains a nucleotide sequence encoding a commercially available embodiment of BCMA CAR, including, for example, idecabtagene vicleucel (ide-cel, also called bb2121). In some embodiments, the polycistronic vector comprises an expression cassette that contains a nucleotide sequence encoding idecabtagene vicleucel or portions thereof. Idecabtagene vicleucel comprises a BCMA CAR with the following components: the BB2121 binder, CD8 α hinge domain, CD8 α transmembrane domain, 4-1BB costimulatory domain, and CD3 ζ signaling domain.

Table M. Exemplary sequences of BCMA CARs

SEQ ID NO:	Sequence	Description
1227	atggccttaccagtaccgccttgctcctgccgtggcctt gctgctccacgccgcccaggccggacatccagatgaccc agtctccatcctccctgtctgcatctgtaggagacagagtc accatcacttgccgggcaagtcagagcattagcagctatt taaattggatcagcagaaaccagggaagcccctaag ctcctgatctatgctgcatccagtttgcaaagtgggtccc atcaagggtcagtggcagtggtatctgggacagatttcactc tcaccatcagcagctctgcaacctgaagattttgcaacttac tactgtcagcaaaaatacgacctcctcacttttgccggag ggaccaaggttgagatcaaaggcagcaccagcggctc cggcaagcctggctctggcgagggcagcacaagggga cagctgcagctgcaggagtcgggcccaggactggtgaa gccttcggagaccctgtccctcacctgcactgtctctggtg gctccatcagcagtagtagttactactggggctggatccg ccagccccaggggaaggggctggagtgattgggagta tctctatagtgaggagcacctactacaacccgtccctcaa gagtcgagtcaccatatccgtagacacgtccaagaacc agttctccctgaagctgagttctgtgaccgccgcagacac ggcgggtgactactgcgccagagatcgtggagacacat actagacgtatggggtcagggtacaatggtcaccgtcag ctcattcgtgcccggtgtcctgcccggccaaacctaccacca cccctgcccctagacctccaccccagccccacaacatcg	Exemplary BCMA CAR nucleotide sequence

SEQ ID NO:	Sequence	Description
	ccagccagcctctgtctctgcggcccgaagcctgtagac ctgctgccggcggagccgtgcacaccagaggcctggac ttgcctgcgacatctacatctgggcccctctggccggca cctgtggcgtgctgctgctgagcctggtgatcacctgtact gcaaccaccggaacaaacggggcagaaagaaactcc tgtatatattcaacaaccatttatgagaccagtacaaact actcaagaggaagatggctgtagctgccgattccagaa gaagaagaaggaggatgtgaactgagagtgaagtca gcagatccgccgacgcccctgcctaccagcagggaca gaaccagctgtacaacgagctgaacctgggcagacgg gaagagtacgacgtgctggacaagcggagaggccgg gaccccgagatgggcggaagcccagacggaagaac cccaggaaggcctgtataacgaactgcagaaagaca agatggccgaggcctacagcgagatcgcatgaagg cgagcggaggcgcggaagggccacgatggcctgtac cagggcctgagcaccgccaccaaggacacctacgacg ccctgcacatgcaggccctgccccccaga	
1228	MALPVTALLPLALLHAARPDIQMTQSP SSLSASVGDRVTITCRASQSISSYLNWYQ QKPGKAPKLLIYAASSLQSGVPSRFSGSG SGTDFTLTISSLQPEDFATYYCQQKYDLL TFGGGTKVEIKGSTSGSGKPGSGEGSTK GQLQLQESGPGLVKPSETLSLTCTVSGG SISSSSYYWGWIRQPPGKGLEWIGSISYS GSTYYNPSLKSRTISVDTSKNQFSLKLS SVTAADTAVYYCARDRGDTILDVWGQGT MTVSSFVPVFLPAKPTTTPAPRPPTPAP TIASQPLSLRPEACRPAAGGAVHTRGLDF ACDIYIWAPLAGTCGVLLLSLVITLYCNHR NKRGRKKLLYIFKQPFMRPVQTTQEEDG CSCRFPEEEEGGCELRVKFSRSADAPAY QQGQNQLYNELNLGRREEYDVLDKRRG RDPEMGGKPRRKNPQEGLYNELQKDKM AEAYSEIGMKGERRRGKGHDGLYQGLST ATKDTYDALHMQALPPR	Exemplary BCMA CAR amino acid sequence

In some embodiments, the antibody portion of the recombinant receptor, e.g., CAR, further includes spacer between the transmembrane domain and extracellular antigen binding domain. In some embodiments, the spacer includes at least a portion of an immunoglobulin constant region, such as a hinge region, e.g., an IgG4 hinge region, and/or a CH1/CL and/or Fc region. In some embodiments, the constant region or portion is of a human IgG, such as IgG4 or IgG1. In some aspects, the portion of the constant region serves as a spacer region between the antigen-

recognition component, e.g., scFv, and transmembrane domain. The spacer can be of a length that provides for increased responsiveness of the cell following antigen binding, as compared to in the absence of the spacer. Exemplary spacers include, but are not limited to, those described in Hudecek et al. (2013) Clin. Cancer Res., 19:3153, WO2014031687, U.S. Patent No. 8,822,647 or published app. No. US 2014/0271635. In some embodiments, the constant region or portion is of a human IgG, such as IgG4 or IgG1.

In some embodiments, the antigen receptor comprises an intracellular domain linked directly or indirectly to the extracellular domain. In some embodiments, the chimeric antigen receptor includes a transmembrane domain linking the extracellular domain and the intracellular signaling domain. In some embodiments, the intracellular signaling domain comprises an ITAM. For example, in some aspects, the antigen recognition domain (e.g. extracellular domain) generally is linked to one or more intracellular signaling components, such as signaling components that mimic activation through an antigen receptor complex, such as a TCR complex, in the case of a CAR, and/or signal via another cell surface receptor. In some embodiments, the chimeric receptor comprises a transmembrane domain linked or fused between the extracellular domain (e.g. scFv) and intracellular signaling domain. Thus, in some embodiments, the antigen-binding component (e.g., antibody) is linked to one or more transmembrane and intracellular signaling domains.

In one embodiment, a transmembrane domain that naturally is associated with one of the domains in the receptor, e.g., CAR, is used. In some instances, the transmembrane domain is selected or modified by amino acid substitution to avoid binding of such domains to the transmembrane domains of the same or different surface membrane proteins to minimize interactions with other members of the receptor complex.

The transmembrane domain in some embodiments is derived either from a natural or from a synthetic source. Where the source is natural, the domain in some aspects is derived from any membrane-bound or transmembrane protein. Transmembrane regions include those derived from (i.e. comprise at least the transmembrane region(s) of) the alpha, beta or zeta chain of the T-cell receptor, CD28, CD3 epsilon, CD45, CD4, CD5, CD8, CD9, CD16, CD22, CD33, CD37, CD64, CD80, CD86, CD 134, CD 137, CD 154. Alternatively, the transmembrane domain in some

embodiments is synthetic. In some aspects, the synthetic transmembrane domain comprises predominantly hydrophobic residues such as leucine and valine. In some aspects, a triplet of phenylalanine, tryptophan and valine will be found at each end of a synthetic transmembrane domain. In some embodiments, the linkage is by linkers, spacers, and/or transmembrane domain(s). In some aspects, the transmembrane domain contains a transmembrane portion of CD28.

In some embodiments, the extracellular domain and transmembrane domain can be linked directly or indirectly. In some embodiments, the extracellular domain and transmembrane are linked by a spacer, such as any described herein. In some embodiments, the receptor contains extracellular portion of the molecule from which the transmembrane domain is derived, such as a CD28 extracellular portion.

Among the intracellular signaling domains are those that mimic or approximate a signal through a natural antigen receptor, a signal through such a receptor in combination with a costimulatory receptor, and/or a signal through a costimulatory receptor alone. In some embodiments, a short oligo- or polypeptide linker, for example, a linker of between 2 and 10 amino acids in length, such as one containing glycines and serines, e.g., glycine-serine doublet, is present and forms a linkage between the transmembrane domain and the cytoplasmic signaling domain of the CAR.

T cell activation is in some aspects described as being mediated by two classes of cytoplasmic signaling sequences: those that initiate antigen-dependent primary activation through the TCR (primary cytoplasmic signaling sequences), and those that act in an antigen-independent manner to provide a secondary or co-stimulatory signal (secondary cytoplasmic signaling sequences). In some aspects, the CAR includes one or both of such signaling components.

The receptor, e.g., the CAR, generally includes at least one intracellular signaling component or components. In some aspects, the CAR includes a primary cytoplasmic signaling sequence that regulates primary activation of the TCR complex. Primary cytoplasmic signaling sequences that act in a stimulatory manner may contain signaling motifs which are known as immunoreceptor tyrosine -based activation motifs or IT AMs. Examples of IT AM containing primary cytoplasmic signaling sequences include those derived from CD3 zeta chain, FcR gamma, CD3 gamma, CD3 delta and CD3 epsilon. In some embodiments, cytoplasmic signaling

molecule(s) in the CAR contain(s) a cytoplasmic signaling domain, portion thereof, or sequence derived from CD3 zeta.

In some embodiments, the receptor includes an intracellular component of a TCR complex, such as a TCR CD3 chain that mediates T-cell activation and cytotoxicity, e.g., CD3 zeta chain. Thus, in some aspects, the antigen-binding portion is linked to one or more cell signaling modules. In some embodiments, cell signaling modules include CD3 transmembrane domain, CD3 intracellular signaling domains, and/or other CD transmembrane domains. In some embodiments, the intracellular component is or includes a CD3-zeta intracellular signaling domain. In some embodiments, the intracellular component is or includes a signaling domain from Fc receptor gamma chain. In some embodiments, the receptor, e.g., CAR, includes the intracellular signaling domain and further includes a portion, such as a transmembrane domain and/or hinge portion, of one or more additional molecules such as CD8, CD4, CD25, or CD 16. For example, in some aspects, the CAR or other chimeric receptor is a chimeric molecule of CD3-zeta (CD3-z) or Fc receptor gamma and a portion of one of CD8, CD4, CD25 or CD16.

In some embodiments, upon ligation of the CAR or other chimeric receptor, the cytoplasmic domain or intracellular signaling domain of the receptor activates at least one of the normal effector functions or responses of the immune cell, e.g., T cell engineered to express the CAR. For example, in some contexts, the CAR induces a function of a T cell such as cytolytic activity or T-helper activity, such as secretion of cytokines or other factors. In some embodiments, a truncated portion of an intracellular signaling domain of an antigen receptor component or costimulatory molecule is used in place of an intact immunostimulatory chain, for example, if it transduces the effector function signal. In some embodiments, the intracellular signaling domain or domains include the cytoplasmic sequences of the T cell receptor (TCR), and in some aspects also those of co-receptors that in the natural context act in concert with such receptors to initiate signal transduction following antigen receptor engagement.

In the context of a natural TCR, full activation generally requires not only signaling through the TCR, but also a costimulatory signal. Thus, in some embodiments, to promote full activation, a component for generating secondary or co-stimulatory signal is also included in the CAR. In other embodiments, the CAR does not include

a component for generating a costimulatory signal. In some aspects, an additional CAR is expressed in the same cell and provides the component for generating the secondary or costimulatory signal.

In some embodiments, the chimeric antigen receptor contains an intracellular domain of a T cell costimulatory molecule. In some embodiments, the CAR includes a signaling domain and/or transmembrane portion of a costimulatory receptor, such as CD28, 4-1BB, OX40, DAP10, and ICOS. In some aspects, the same CAR includes both the activating and costimulatory components. In some embodiments, the chimeric antigen receptor contains an intracellular domain derived from a T cell costimulatory molecule or a functional variant thereof, such as between the transmembrane domain and intracellular signaling domain. In some aspects, the T cell costimulatory molecule is CD28 or 41BB.

In some embodiments, the activating domain is included within one CAR, whereas the costimulatory component is provided by another CAR recognizing another antigen. In some embodiments, the CARs include activating or stimulatory CARs, costimulatory CARs, both expressed on the same cell (see WO2014/055668). In some aspects, the cells include one or more stimulatory or activating CAR and/or a costimulatory CAR. In some embodiments, the cells further include inhibitory CARs (iCARs, see Fedorov et al., *Sci. Transl. Medicine*, 5(215) (December, 2013), such as a CAR recognizing an antigen other than the one associated with and/or specific for the disease or condition whereby an activating signal delivered through the disease-targeting CAR is diminished or inhibited by binding of the inhibitory CAR to its ligand, e.g., to reduce off-target effects.

In certain embodiments, the intracellular signaling domain comprises a CD28 transmembrane and signaling domain linked to a CD3 (e.g., CD3-zeta) intracellular domain. In some embodiments, the intracellular signaling domain comprises a chimeric CD28 and CD137 (4-1BB, TNFRSF9) co-stimulatory domains, linked to a CD3 zeta intracellular domain.

In some embodiments, the CAR encompasses one or more, e.g., two or more, costimulatory domains and an activation domain, e.g., primary activation domain, in the cytoplasmic portion. Exemplary CARs include intracellular components of CD3-zeta, CD28, and 4-1BB.

In some embodiments the intracellular signaling domain includes intracellular components of a 4-1BB signaling domain and a CD3-zeta signaling domain. In some embodiments, the intracellular signaling domain includes intracellular components of a CD28 signaling domain and a CD3zeta signaling domain. In some embodiments, the CAR comprises an extracellular antigen binding domain (e.g., antibody or antibody fragment, such as an scFv) that binds to an antigen (e.g. tumor antigen), a spacer (e.g. containing a hinge domain, such as any as described herein), a transmembrane domain (e.g. any as described herein), and an intracellular signaling domain (e.g. any intracellular signaling domain, such as a primary signaling domain or costimulatory signaling domain as described herein). In some embodiments, the intracellular signaling domain is or includes a primary cytoplasmic signaling domain. In some embodiments, the intracellular signaling domain additionally includes an intracellular signaling domain of a costimulatory molecule (e.g., a costimulatory domain). Examples of exemplary components of a CAR are described in **Table 14**. In provided aspects, the sequences of each component in a CAR can include any combination listed in **Table 14**.

Table 14

Component	Sequence	SEQ ID NO:
<i>Extracellular binding domain</i>		
Anti-CD19 scFv (FMC63)	DIQMTQTTSSLSASLGDRVTISCRASQDISKYLNWYQQKPD GTVKLLIYHTSRLHSGVPSRFSGSGSGTDYSLTISNLEQEDI ATYFCQQGNTLPYTFGGGKLEITGSTSGSGKPGSGEGST KGEVKLQESGPGLVAPSQSLSVTCTVSGVSLPDYGVSWIR QPPRKGLEWLGVIWGSETTYNSALKSRLTIKDNSKSQVF LKMNSLQTDDTAIYYCAKHYYYGGSYAMDYWGQGTSTVTV SS	1075
Anti-CD19 scFv (FMC63)	DIQMTQTTSSLSASLGDRVTISCRASQDISKYLNWYQQKPD GTVKLLIYHTSRLHSGVPSRFSGSGSGTDYSLTISNLEQEDI ATYFCQQGNTLPYTFGGGKLEITGGGGSGGGGSGGGGS EVKLQESGPGLVAPSQSLSVTCTVSGVSLPDYGVSWIRQP PRKGLEWLGVIWGSETTYNSALKSRLTIKDNSKSQVFLK MNSLQTDDTAIYYCAKHYYYGGSYAMDYWGQGTSTVTVSS	1076
Anti-BCMA sdAb (FHVH74)	QVQLVESGGGLVQPGGSLRLSCAASGFTFTNHAMSWVRQ APGKGLELVSSISGNRTTYADSVKGRFTISRDISKNTLDL QMNSLRAEDTAVYYCAKDGGGETLVDSRGQGTSLVTVSS	1077

Anti-BCMA sdAb (FHVH32)	QVQLVESGGGLVQPGGSLRLSCAASGFTFSSHAMTWVRQ APGKGLEWVAISGSGDFTHYADSVKGRFTISRDNKNTV SLQMNNLRAEDTAVYYCAKDEDGGSLLGYRGQGTTLTVS S	1078
Anti-BCMA sdAb (FHVH33)	EVQLLES GGGLVQP GGSLRLSCAASGFTFSSYAMSWVRQ APGKGLEWVSSISGSGDYIYYADSVKGRFTISRDNKNTLYL QMNSLRAEDTAVYYCAKEGTGANSSLADYRGQGTTLTVS S	1079
Anti-BCMA sdAb (FHVH93)	EVQLLES GGGLIQP GGSLRLSCAASGFTFSSHAMTWVRQA PGKGLEWWSAISGSGDYTHYADSVKGRFTISRDNKNTVY LQMNSLRAEDSAVYYCAKDEDGGSLLGHRGQGTTLTVSS	1080
<i>Spacer (e.g. hinge)</i>		
IgG4 Hinge	ESKYGPPCPPCP	1081
CD8 Hinge	TTTPAPRPPTPAPTIASQPLSLRPE	1082
CD28	IEVMYPPPYLDNEKSNGTIIHVKGKHLCPSPFPGPSKP	1083
<i>Transmembrane</i>		
CD8	ACRPAAGGAVHTRGLDFACDIYWAPLAGTCGVLLLSLVITL YC	1084
CD28	FWVLVVVGGLVACYSLLVTVAFIIFWW	1085
CD28	MFWVLVVVGGLVACYSLLVTVAFIIFWW	1214
<i>Costimulatory domain</i>		
CD28	RSKRSRLHSDYMNMTPRRPGPTRKHYPYAPPRDFAAY RS	1086
4-1BB	KRGRKKLLYIFKQPFMRPVQTTQEEDGCSCRFEEEEGGC EL	1087
<i>Primary Signaling Domain</i>		
CD3zeta	RVKFSRSADAPAYQQGQNQLYNELNLGRREEYDVLDKRR GRDPEMGGKPRRKNPQEGLYNELQKDKMAEAYSEIGMKG ERRRGKGHDGLYQGLSTATKDTYDALHMQALPPR	1088
CD3zeta (Q>K)	RVKFSRSADAPAYKQGQNQLYNELNLGRREEYDVLDKRR GRDPEMGGKPRRKNPQEGLYNELQKDKMAEAYSEIGMKG ERRRGKGHDGLYQGLSTATKDTYDALHMQALPPR	1089

In some embodiments, the antigen receptor further includes a marker and/or cells expressing the CAR or other antigen receptor further includes a surrogate marker, such as a cell surface marker, which may be used to confirm transduction or

engineering of the cell to express the receptor. In some aspects, the marker includes all or part (e.g., truncated form) of CD34, a NGFR, or epidermal growth factor receptor, such as truncated version of such a cell surface receptor (e.g., tEGFR). In some embodiments, the nucleic acid encoding the marker is operably linked to a polynucleotide encoding for a linker sequence, such as a cleavable linker sequence, e.g., T2A. For example, a marker, and optionally a linker sequence, can be any as disclosed in published patent application No. WO2014031687. For example, the marker can be a truncated EGFR (tEGFR) that is, optionally, linked to a linker sequence, such as a T2A cleavable linker sequence.

In some embodiments, the marker is a molecule, e.g., cell surface protein, not naturally found on T cells or not naturally found on the surface of T cells, or a portion thereof. In some embodiments, the molecule is a non-self molecule, e.g., non-self protein, i.e., one that is not recognized as "self" by the immune system of the host into which the cells will be adoptively transferred.

In some embodiments, the marker serves no therapeutic function and/or produces no effect other than to be used as a marker for genetic engineering, e.g., for selecting cells successfully engineered. In other embodiments, the marker may be a therapeutic molecule or molecule otherwise exerting some desired effect, such as a ligand for a cell to be encountered in vivo, such as a costimulatory or immune checkpoint molecule to enhance and/or dampen responses of the cells upon adoptive transfer and encounter with ligand.

In some cases, CARs are referred to as first, second, and/or third generation CARs. In some aspects, a first generation CAR is one that solely provides a CD3-chain induced signal upon antigen binding; in some aspects, a second-generation CARs is one that provides such a signal and costimulatory signal, such as one including an intracellular signaling domain from a costimulatory receptor such as CD28 or CD 137; in some aspects, a third generation CAR is one that includes multiple costimulatory domains of different costimulatory receptors.

For example, in some embodiments, the CAR contains an antibody, e.g., an antibody fragment, a transmembrane domain that is or contains a transmembrane portion of CD28 or a functional variant thereof, and an intracellular signaling domain containing a signaling portion of CD28 or functional variant thereof and a signaling portion of CD3 zeta or functional variant thereof. In some embodiments, the CAR

contains an antibody, e.g., antibody fragment, a transmembrane domain that is or contains a transmembrane portion of CD28 or a functional variant thereof, and an intracellular signaling domain containing a signaling portion of a 4- IBB or functional variant thereof and a signaling portion of CD3 zeta or functional variant thereof. In some such embodiments, the receptor further includes a spacer containing a portion of an Ig molecule, such as a human Ig molecule, such as an Ig hinge, e.g. an IgG4 hinge, such as a hinge -only spacer.

In some aspects, the spacer contains only a hinge region of an IgG, such as only a hinge of IgG4 or IgG. In other embodiments, the spacer is or contains an Ig hinge, e.g., an IgG4-derived hinge, optionally linked to a CH2 and/or CH3 domains. In some embodiments, the spacer is an Ig hinge, e.g., an IgG4 hinge, linked to CH2 and CH3 domains. In some embodiments, the spacer is an Ig hinge, e.g., an IgG4 hinge, linked to a CH3 domain only. In some embodiments, the spacer is or comprises a glycine-serine rich sequence or other flexible linker such as known flexible linkers. For example, in some embodiments, the CAR includes an antibody such as an antibody fragment, including scFvs, a spacer, such as a spacer containing a portion of an immunoglobulin molecule, such as a hinge region and/or one or more constant regions of a heavy chain molecule, such as an Ig-hinge containing spacer, a transmembrane domain containing all or a portion of a CD28-derived transmembrane domain, a CD28 -derived intracellular signaling domain, and a CD3 zeta signaling domain. In some embodiments, the CAR includes an antibody or fragment, such as scFv, a spacer such as any of the Ig-hinge containing spacers, a CD28-derived transmembrane domain, a 4-IBB-derived intracellular signaling domain, and a CD3 zeta-derived signaling domain.

The recombinant receptors, such as CARs, expressed by the cells administered to the subject generally recognize or specifically bind to a molecule that is expressed in, associated with, and/or specific for the disease or condition or cells thereof being treated. Upon specific binding to the molecule, e.g., antigen, the receptor generally delivers an immunostimulatory signal, such as an ITAM-transduced signal, into the cell, thereby promoting an immune response targeted to the disease or condition. For example, in some embodiments, the cells express a CAR that specifically binds to an antigen expressed by a cell or tissue of the disease or condition or associated with the disease or condition.

b. *T Cell Receptors Antigen Receptors (TCRs)*

In some embodiments, engineered cells, such as T cells, used in connection with the provided methods, uses, articles of manufacture or compositions are cells that express a T cell receptor (TCR) or antigen-binding portion thereof that recognizes an peptide epitope or T cell epitope of a target polypeptide, such as an antigen of a tumor, viral or autoimmune protein.

In some embodiments, a “T cell receptor” or “TCR” is a molecule that contains a variable a and b chains (also known as TCRalpha and TCRbeta, respectively) or a variable g and d chains (also known as TCRalpha and TCRbeta, respectively), or antigen-binding portions thereof, and which is capable of specifically binding to a peptide bound to an MHC molecule. In some embodiments, the TCR is in the ab form. Typically, TCRs that exist in ab and gd forms are generally structurally similar, but T cells expressing them may have distinct anatomical locations or functions. A TCR can be found on the surface of a cell or in soluble form. Generally, a TCR is found on the surface of T cells (or T lymphocytes) where it is generally responsible for recognizing antigens bound to major histocompatibility complex (MHC) molecules.

Unless otherwise stated, the term “TCR” should be understood to encompass full TCRs as well as antigen-binding portions or antigen-binding fragments thereof. In some embodiments, the TCR is an intact or full-length TCR, including TCRs in the ab form or gd form. In some embodiments, the TCR is an antigen-binding portion that is less than a full-length TCR but that binds to a specific peptide bound in an MHC molecule, such as binds to an MHC-peptide complex. In some cases, an antigen-binding portion or fragment of a TCR can contain only a portion of the structural domains of a full-length or intact TCR, but yet is able to bind the peptide epitope, such as MHC-peptide complex, to which the full TCR binds. In some cases, an antigen-binding portion contains the variable domains of a TCR, such as variable a chain and variable b chain of a TCR, sufficient to form a binding site for binding to a specific MHC-peptide complex. Generally, the variable chains of a TCR contain complementarity determining regions involved in recognition of the peptide, MHC and/or MHC-peptide complex.

c. *Multi-Targeting*

In some embodiments, the cells used in connection with the provided methods, uses, articles of manufacture and compositions include cells employing multi-targeting strategies, such as expression of two or more genetically engineered receptors on the cell, each recognizing the same or a different antigen and typically each including a different intracellular signaling component. Such multi-targeting strategies are described, for example, in WO 2014055668 (describing combinations of activating and costimulatory CARs, e.g., targeting two different antigens present individually on off-target, e.g., normal cells, but present together only on cells of the disease or condition to be treated) and Fedorov et al., *Sci. Transl. Medicine*, 5(215) (2013) (describing cells expressing an activating and an inhibitory CAR, such as those in which the activating CAR binds to one antigen expressed on both normal or non-diseased cells and cells of the disease or condition to be treated, and the inhibitory CAR binds to another antigen expressed only on the normal cells or cells which it is not desired to treat).

For example, in some embodiments, the cells include a receptor expressing a first genetically engineered antigen receptor (e.g., CAR) which is capable of inducing an activating or stimulatory signal to the cell, generally upon specific binding to the antigen recognized by the first receptor, e.g., the first antigen. In some embodiments, the cell further includes a second genetically engineered antigen receptor (e.g., CAR), e.g., a chimeric costimulatory receptor, which is capable of inducing a costimulatory signal to the immune cell, generally upon specific binding to a second antigen recognized by the second receptor. In some embodiments, the first antigen and second antigen are the same. In some embodiments, the first antigen and second antigen are different.

In some embodiments, the first and/or second genetically engineered antigen receptor (e.g. CAR) is capable of inducing an activating signal to the cell. In some embodiments, the receptor includes an intracellular signaling component containing ITAM or ITAM-like motifs. In some embodiments, the activation induced by the first receptor involves a signal transduction or change in protein expression in the cell resulting in initiation of an immune response, such as ITAM phosphorylation and/or initiation of ITAM-mediated signal transduction cascade, formation of an

immunological synapse and/or clustering of molecules near the bound receptor (e.g. CD4 or CD8, etc.), activation of one or more transcription factors, such as NF-KB and/or AP-1, and/or induction of gene expression of factors such as cytokines, proliferation, and/or survival.

In some embodiments, the first and/or second receptor includes intracellular signaling domains or regions of costimulatory receptors such as CD28, CD137 (4-1BB), OX40, and/or ICOS. In some embodiments, the first and second receptor include an intracellular signaling domain of a costimulatory receptor that are different. In one embodiment, the first receptor contains a CD28 costimulatory signaling region and the second receptor contain a 4-1BB co-stimulatory signaling region or vice versa.

In some embodiments, the first and/or second receptor includes both an intracellular signaling domain containing ITAM or ITAM-like motifs and an intracellular signaling domain of a costimulatory receptor.

In some embodiments, the first receptor contains an intracellular signaling domain containing ITAM or ITAM-like motifs and the second receptor contains an intracellular signaling domain of a costimulatory receptor. The costimulatory signal in combination with the activating signal induced in the same cell is one that results in an immune response, such as a robust and sustained immune response, such as increased gene expression, secretion of cytokines and other factors, and T cell mediated effector functions such as cell killing.

In some embodiments, neither ligation of the first receptor alone nor ligation of the second receptor alone induces a robust immune response. In some aspects, if only one receptor is ligated, the cell becomes tolerized or unresponsive to antigen, or inhibited, and/or is not induced to proliferate or secrete factors or carry out effector functions. In some such embodiments, however, when the plurality of receptors are ligated, such as upon encounter of a cell expressing the first and second antigens, a desired response is achieved, such as full immune activation or stimulation, e.g., as indicated by secretion of one or more cytokine, proliferation, persistence, and/or carrying out an immune effector function such as cytotoxic killing of a target cell.

In some embodiments, the two receptors induce, respectively, an activating and an inhibitory signal to the cell, such that binding by one of the receptor to its antigen activates the cell or induces a response, but binding by the second inhibitory

receptor to its antigen induces a signal that suppresses or dampens that response. Examples are combinations of activating CARs and inhibitory CARs or iCARs. Such a strategy may be used, for example, in which the activating CAR binds an antigen expressed in a disease or condition but which is also expressed on normal cells, and the inhibitory receptor binds to a separate antigen which is expressed on the normal cells but not cells of the disease or condition.

In some embodiments, the multi-targeting strategy is employed in a case where an antigen associated with a particular disease or condition is expressed on a non-diseased cell and/or is expressed on the engineered cell itself, either transiently (e.g., upon stimulation in association with genetic engineering) or permanently. In such cases, by requiring ligation of two separate and individually specific antigen receptors, specificity, selectivity, and/or efficacy may be improved.

In some embodiments, the plurality of antigens, e.g., the first and second antigens, are expressed on the cell, tissue, or disease or condition being targeted, such as on the cancer cell. In some aspects, the cell, tissue, disease or condition is multiple myeloma or a multiple myeloma cell. In some embodiments, one or more of the plurality of antigens generally also is expressed on a cell which it is not desired to target with the cell therapy, such as a normal or non-diseased cell or tissue, and/or the engineered cells themselves. In such embodiments, by requiring ligation of multiple receptors to achieve a response of the cell, specificity and/or efficacy is achieved.

d. *Chimeric Auto-Antibody Receptor (CAAR)*

In some embodiments, the recombinant receptor is a chimeric autoantibody receptor (CAAR). In some embodiments, the CAAR binds, e.g., specifically binds, or recognizes, an autoantibody. In some embodiments, a cell expressing the CAAR, such as a T cell engineered to express a CAAR, can be used to bind to and kill autoantibody-expressing cells, but not normal antibody expressing cells. In some embodiments, CAAR-expressing cells can be used to treat an autoimmune disease associated with expression of self-antigens, such as autoimmune diseases. In some embodiments, CAAR-expressing cells can target B cells that ultimately produce the autoantibodies and display the autoantibodies on their cell surfaces, mark these B cells as disease-specific targets for therapeutic intervention. In some embodiments,

CAAR-expressing cells can be used to efficiently targeting and killing the pathogenic B cells in autoimmune diseases by targeting the disease-causing B cells using an antigen-specific chimeric autoantibody receptor. In some embodiments, the recombinant receptor is a CAAR, such as any described in U.S. Patent Application Pub. No. US 2017/0051035.

In some embodiments, the CAAR comprises an autoantibody binding domain, a transmembrane domain, and one or more intracellular signaling region or domain (also interchangeably called a cytoplasmic signaling domain or region). In some embodiments, the intracellular signaling region comprises an intracellular signaling domain. In some embodiments, the intracellular signaling domain is or comprises a primary signaling domain, a signaling domain that is capable of stimulating and/or inducing a primary activation signal in a T cell, a signaling domain of a T cell receptor (TCR) component (e.g. an intracellular signaling domain or region of a CD3-zeta) chain or a functional variant or signaling portion thereof), and/or a signaling domain comprising an immunoreceptor tyrosine- based activation motif (ITAM).

In some embodiments, the autoantibody binding domain comprises an autoantigen or a fragment thereof. The choice of autoantigen can depend upon the type of autoantibody being targeted. For example, the autoantigen may be chosen because it recognizes an autoantibody on a target cell, such as a B cell, associated with a particular disease state, e.g. an autoimmune disease, such as an autoantibody-mediated autoimmune disease. In some embodiments, the autoimmune disease includes pemphigus vulgaris (PV). Exemplary autoantigens include desmoglein 1 (Dsg1) and Dsg3.

In some embodiments, the encoded nucleic acid is operatively linked to a “positive target cell-specific regulatory element” (or positive TCSRE). In some embodiments, the positive TCSRE is a functional nucleic acid sequence. In some embodiments, the positive TCSRE comprises a promoter or enhancer. In some embodiments, the TCSRE is a nucleic acid sequence that increases the level of an exogenous agent in a target cell. In some embodiments, the positive target cell-specific regulatory element comprises a T cell-specific promoter, a T cell-specific enhancer, a T cell-specific splice site, a T cell-specific site extending half-life of an RNA or protein, a T cell-specific mRNA nuclear export promoting site, a T cell-specific translational enhancing site, or a T cell-specific post-translational modification site. In some

embodiments, the T cell-specific promoter is a promoter described in Immgen consortium, herein incorporated by reference in its entirety, e.g., the T cell-specific promoter is an IL2RA (CD25), LRR32, FOXP3, or IKZF2 promoter. In some embodiments, the T cell-specific promoter or enhancer is a promoter or enhancer described in Schmidl et al., Blood. 2014 Apr 24;123(17):e68-78., herein incorporated by reference in its entirety. In some embodiments, the T cell-specific promoter is a transcriptionally active fragment of any of the foregoing. In some embodiments, the T-cell specific promoter is a variant having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity to any of the foregoing.

In some embodiments, the encoded nucleic acid is operatively linked to a “negative target cell-specific regulatory element” (or negative TCSRE). In some embodiments, the negative TCSRE is a functional nucleic acid sequence. In some embodiments, the negative TCSRE is a miRNA recognition site that causes degradation or inhibition of the viral vector in a non-target cell. In some embodiments, the exogenous agent is operatively linked to a “non-target cell-specific regulatory element” (or NTCSRE). In some embodiments, the NTCSRE comprises a nucleic acid sequence that decreases the level of an exogenous agent in a non-target cell compared to in a target cell. In some embodiments, the NTCSRE comprises a non-target cell-specific miRNA recognition sequence, non-target cell-specific protease recognition site, non-target cell-specific ubiquitin ligase site, non-target cell-specific transcriptional repression site, or non-target cell-specific epigenetic repression site. In some embodiments, the NTCSRE comprises a tissue-specific miRNA recognition sequence, tissue-specific protease recognition site, tissue-specific ubiquitin ligase site, tissue-specific transcriptional repression site, or tissue-specific epigenetic repression site. In some embodiments, the NTCSRE comprises a non-target cell-specific miRNA recognition sequence, non-target cell-specific protease recognition site, non-target cell-specific ubiquitin ligase site, non-target cell-specific transcriptional repression site, or non-target cell-specific epigenetic repression site. In some embodiments, the NTCSRE comprises a non-target cell-specific miRNA recognition sequence and the miRNA recognition sequence is able to be bound by one or more of miR31, miR363, or miR29c. In some embodiments, the NTCSRE is situated or encoded within a transcribed region encoding the exogenous agent,

optionally wherein an RNA produced by the transcribed region comprises the miRNA recognition sequence within a UTR or coding region.

In some embodiments, the viral vector comprising an anti-CD8 scFv or sdAb composition described herein can be administered to a subject, e.g., a mammal, e.g., a human. In such embodiments, the subject may be at risk of, may have a symptom of, or may be diagnosed with or identified as having, a particular disease or condition (e.g., a disease or condition described herein).

In some aspects, resting or non-activated T cells are contacted with a viral vector of the disclosure (e.g., a retroviral vector or lentiviral vector) that includes a CD8 binding agent. The contacting may be performed in vitro (e.g., with T cells derived from a healthy donor or a donor in need of cellular therapy) or in vivo by administration of the viral vector to a subject.

In some embodiments, the resting or non-activated T cells are not treated with one or more T cell stimulatory molecules (e.g., an anti CD-3 antibody), one or more T cell costimulatory molecules, and/or one or more T cell activating cytokines. In some embodiments, the resting or non-activated T cells are not treated with any of one or more T cell stimulatory molecules (e.g., an anti CD-3 antibody), one or more T cell costimulatory molecules, and/or one or more T cell activating cytokines.

In additional aspects, the application includes methods of administration to a subject of a viral vector that includes an anti-CD8 binding agent, wherein the subject is not administered or has not been administered a T cell activating treatment. In some embodiments, the T cell activating treatment includes one or more T cell stimulatory molecules (e.g., an anti CD-3 antibody), one or more T cell costimulatory molecules, and/or one or more T cell activating cytokines. In some embodiments, the subject is not administered or has not been administered any of one or more T cell stimulatory molecules (e.g., an anti CD-3 antibody), one or more T cell costimulatory molecules, and/or one or more T cell activating cytokines. In some embodiments, the T cell activating treatment is lymphodepletion. In certain embodiments, the subject is not administered or has not been administered the T cell activating treatment within 1 month before or after administration of the viral vector. In some embodiments, the subject is not administered or has not been administered the T cell activating treatment within 1 month before administration of the viral vector, such as within or at or about 4 weeks, 3 weeks, 2 weeks or 1 weeks, such as at or about 1 day, 2 days, 3

days, 4 days, 5 days, 6 days or 7 days before administration of the viral vector. In some embodiments, the subject is not administered the T cell activating treatment within 1 month after administration of the viral vector, such as within or at or about 4 weeks, 3 weeks, 2 weeks or 1 weeks, such as at or about 1 day, 2 days, 3 days, 4 days, 5 days, 6 days or 7 days after administration of the viral vector.

In some aspects, the viral vectors of the disclosure do not include one or more T cell stimulatory molecules (e.g., an anti CD-3 antibody), one or more T cell costimulatory molecules, and/or one or more T cell activating cytokines.

The use of anti-CD3 antibodies is well-known for activation of T cells. The anti-CD3 antibodies can be of any species, e.g., mouse, rabbit, human, humanized, or camelid. Exemplary antibodies include OKT3, CRIS-7, I2C the anti-CD3 antibody included in DYNABEADS Human T-Activator CD3/CD28 (Thermo Fisher), and the anti-CD3 domains of approved and clinically studied molecules such as blinatumomab, catumaxomab, fotetuzumab, teclistamab, ertumaxomab, epcoritamab, talquetamab, odronextamab, cibistamab, obrindatamab, tidutamab, duvortuxizumab, solitomab, eluvixtamab, pavurutamab, tepoditamab, vibecotamab, plamotamab, glofitamab, etevritamab, and tarlatamab.

In some embodiments, the one or more T cell costimulatory molecules include CD28 ligands (e.g., CD80 and CD86); antibodies that bind to CD28 such as CD28.2, the anti-CD28 antibody included in DYNABEADS Human T-Activator CD3/CD28 (Thermo Fisher) and anti-CD28 domains disclosed in US2020/0199234, US2020/0223925, US2020/0181260, US2020/0239576, US2020/0199233, US2019/0389951, US2020/0299388, US2020/0399369, and US2020/0140552; CD137 ligand (CD137L); anti-CD137 antibodies such as urelumab and utomilumab; ICOS ligand (ICOS-L); and anti-ICOS antibodies such as feladilimab, vopratelimab, and the anti-ICOS domain of izuralimab.

In some embodiments, the one or more T cell activating cytokines include IL-2, IL-7, IL-15, IL-21, interferons (e.g., interferon-gamma), and functional variants and modified versions thereof.

Lymphodepletion may be induced by various treatments that destroy lymphocytes and T cells in the subject. For example, the lymphodepletion may include myeloablative chemotherapies, such as fludarabine, cyclophosphamide,

bendamustine, and combinations thereof. Lymphodepletion may also be induced by irradiation (e.g., full-body irradiation) of the subject.

In some embodiments, the source of targeted lipid particles are from the same subject that is administered a targeted viral vector composition. In other embodiments, they are different. In some embodiments, the source of targeted viral vectors and recipient tissue may be autologous (from the same subject) or heterologous (from different subjects). In some embodiments, the donor tissue for targeted viral vector compositions described herein may be a different tissue type than the recipient tissue. In some embodiments, the donor tissue may be muscular tissue and the recipient tissue may be connective tissue (e.g., adipose tissue). In other embodiments, the donor tissue and recipient tissue may be of the same or different type, but from different organ systems.

In some embodiments, the targeted lipid particles (e.g., viral vector) composition described herein may be administered to a subject having a cancer, an autoimmune disease, an infectious disease, a metabolic disease, a neurodegenerative disease, or a genetic disease (e.g., enzyme deficiency). In some embodiments, the subject is in need of regeneration.

In some embodiments, the cancer is a T cell-mediated cancer. In another embodiment, the antigen binding moiety portion of a CAR is designed to treat a particular cancer. In some embodiments, the targeted viral vector can be used to treat cancers and disorders including but are not limited to Non-Hodgkin lymphoma (NHL), acute lymphocytic leukemia (ALL), chronic lymphocytic leukemia (CLL), multiple myeloma, and the like. In some embodiments, the targeted viral vector can be used to treat B cell malignancies, e.g., refractory B cell malignancies.

In some embodiments, the targeted viral vector is co-administered with an inhibitor of a protein that inhibits membrane fusion. For example, Suppressyn is a human protein that inhibits cell-cell fusion (Sugimoto et al., "A novel human endogenous retroviral protein inhibits cell-cell fusion" Scientific Reports 3: 1462 DOI: 10.1038/srep01462). In some embodiments, the targeted lipid particle is co-administered with an inhibitor of suppressyn, e.g., a siRNA or inhibitory antibody.

Examples

The present disclosure may be further described by the following non-limiting examples, in which standard techniques known to the skilled artisan and techniques analogous to those described in these examples may be used where appropriate. It is understood that the skilled artisan will envision additional embodiments consistent with the disclosure provided herein.

Example 1: Characterization of CD8 scFvs and VHH

This example describes methods to generate and characterize functional titers of Nipah G/F pseudo-typed lentivirus functionalized with a CD8-specific scFv or VHH. Eliminating binding of Nipah G to EphrinB2/B3 and functionalizing Nipah G with a CD8-specific binder allows for targeting of CD8-expressing cells with the lentiviral vector.

Lentiviral production was performed as follows: HEK-293LX cells were plated 24 hours in advance of transfection. On the day of transfection, HEK-293LX cells were transfected with a lentiviral packaging plasmid, a lentiviral transfer plasmid encoding GFP (pSFFV-GFP), and plasmids encoding for Nipah G protein retargeted for CD8 receptor targeting (NiV-G(CD8)) and Nipah F fusion protein (NiV-Fd22). After 24 hours, a media change was performed, and after another 24 hours, the lentivirus was harvested. To harvest the lentivirus, supernatant was removed from the HEK293LX cells and spun at 1000xg for 5 minutes. The supernatant was removed and immediately added to CD8-positive target cells or T cells, or frozen at -80 °C for later use.

Several cell lines were used to characterize specificity and transduction efficiency of the lentiviral vectors described above. A SupT1 CD8 $\alpha\beta$ knockout line was generated from SupT1 human T lymphoblast cells. On the day of transduction, SupT1 and SupT1 CD8 $\alpha\beta$ knock out cells, HEK293LX cells expressing M. nemestrina CD8 $\alpha\beta$ and HEK293LX background lines were plated in a 96 well cell plate. Two hours later, the lentivirus was serially diluted and added to the recipient cells. The lentivirus and cells were incubated at 37°C at 5% CO₂ for three days and analyzed by flow cytometry for GFP expression. Briefly, HEK293LX and HEK293LX cells overexpressing M. nemestrina CD8 $\alpha\beta$ cells were harvested from the 96-well cell plate by trypsinization, transferred to a 96-well U-bottom sample plate, and pelleted

by centrifugation at 1000 x g for 5 min. The cells were then resuspended in 200 μ L of PBS before flow cytometry analysis. SupT1 and SupT1 CD8 $\alpha\beta$ knock-out cells were pelleted by centrifugation and resuspended in 200 μ L PBS + 2% FBS for flow cytometry analysis.

The cells were measured for GFP fluorescence using a BD Celesta cytometer. GFP was excited with a 488nm laser and emission captured at 513 $\pm$ 26nm. Forward and side scatter gating was initially used to capture cell-sized events and discard small debris. Events positive for GFP were determined by gating at the minimum level for which the negative control cell samples (cells not treated with lentivirus) showed <0.5% of events positive for GFP expression. The gated cells positive for GFP fluorescence were then assessed for the % of GFP-positive cells of the total cells. To calculate lentiviral functional titer, a transduced cell well showing a GFP % positive that was between 5% and 20% of cells was used to determine titer. The formula for virus titer calculation: $\text{titer} = \{(F \times C_n) / V\} \times DF$. F is the frequency of GFP-positive cells determined by flow cytometry; C_n = The total number of target cells infected. V = The volume of the inoculum, and DF = dilution factor (see Fig. 1).

Lentivirus with CD8 binders 97 (VHH) and 46 (scFv) were further tested on additional non-CD8-expressing cell lines and the CD8 positive cell line SupT1 as above (see Fig. 2). Maximal transduction of the cells over various dilution factors is indicated below (**Table 15**). VSV-G was used as a non-cell specific control. These results indicate that the binders are specific for CD8.

Table 15.

	CD8 Binder 97	CD8 Binder 46	VSV-G
SupT1	87.0	99.0	98.0
293 Lenti-X	6.0	3.0	81.0
Ramos	9.0	4.0	68.0
HUVEC	10.0	5.0	94.0
C2C12	0	0	63.0
HEL 92.1.7	0	2.0	100.0
HeLa	1.0	0	97.0
HepG2	0	0	99.0
hFOB	0.1	0.6	68.3
HULEC-5a	2.0	4.0	83.0
Kasumi 1	0	3.0	74.0
SK-N-AS	1.0	2.0	99.0
U-937	0	0	100.0

Binding to human and *M. nemestrina* CD8 was also evaluated. On the day of transduction, SupT1 and SupT1 CD8 $\alpha\beta$ knockout cell lines, HEK293LX cells expressing *M. nemestrina* CD8 $\alpha\beta$ and HEK293LX background lines were plated in a 96 well cell plate. Two hours later, the lentivirus was serially diluted and added to the recipient cells. The lentivirus and cells were incubated at 37°C at 5% CO₂ for three days and analyzed by flow cytometry for GFP expression. Briefly, HEK293LX and HEK293LX cells overexpressing *M. nemestrina* CD8 $\alpha\beta$ cells were harvested from the 96-well cell plate by trypsinization, transferred to a 96-well U-bottom sample plate, and pelleted by centrifugation at 1000 x g for 5 min. The cells were then resuspended in 200 μ L of PBS before flow cytometry analysis. SupT1 and SupT1 CD8 $\alpha\beta$ knock-out cells were pelleted by centrifugation and resuspended in 200 μ L PBS + 2% FBS for flow cytometry analysis.

The cells were measured for GFP fluorescence using a BD Celesta cytometer. GFP was excited with a 488nm laser and emission captured at 513 $\pm$ 26nm. Forward and side scatter gating was initially used to capture cell-sized events and discard small debris. Events positive for GFP were determined by gating at the minimum level for which the negative control cell samples (cells not treated with lentivirus) showed <0.5% of events positive for GFP expression. The gated cells positive for GFP fluorescence were then assessed for the % of GFP-positive cells of the total cells. To calculate lentiviral functional titer, a transduced cell well showing a GFP % positive that was between 5% and 20% of cells was used to determine titer. The formula for virus titer calculation: $\text{titer} = \{(F \times C_n) / V\} \times DF$. F is the frequency of GFP-positive cells determined by flow cytometry; C_n = The total number of target cells infected. V = The volume of the inoculum, and DF = dilution factor. Titer for human CD8 (SupT1) is presented in Fig. 2. Titer for *M. nemestrina* CD8 (HEK293LX cells expressing *M. nemestrina* CD8 $\alpha\beta$) is presented in Fig. 3.

Lentiviral production was performed as follows: LV-Max HEK293 cells were seeded 24 hours in advance of transfection. On the day of transfection, LV-Max HEK293 cells were transfected with lentiviral packaging plasmid, the lentiviral transfer plasmid encoding GFP (pSFFV-GFP), and plasmids encoding Nipah G protein retargeted for CD8 receptor targeting (NiV-G(CD8)) and Nipah F fusion protein (NiV-Fd22). Two days later supernatant from the transfected cells was harvested, passed through a 0.45 μ m filter, and concentrated by ultracentrifugation at 120,000 x g for 90 minutes.

After ultracentrifugation the lentiviral pellet was resuspended in PBS, and either used immediately for transduction or aliquoted and frozen at -80°C for use later.

Human CD8 α and human CD8 $\alpha\beta$ overexpressing cell lines were generated from HEK293LX cells. On the day of transduction, human CD8 α , human CD8 $\alpha\beta$ overexpressing and HEK293LX background cell lines were plated in a 96 well cell plate. Two hours later, the lentivirus was serially diluted and added to the recipient cells. The lentivirus and cells were incubated at 37°C at 5% CO₂ for three days and analyzed by flow cytometry for GFP expression. Briefly, human CD8 α , human CD8 $\alpha\beta$ overexpressing, and HEK293LX background cell lines were harvested from the 96-well cell plate by trypsinization, transferred to a 96-well U-bottom sample plate, and pelleted by centrifugation at 1000 x g for 5 min. The cells were then resuspended in 200 μ L of PBS before flow cytometry analysis.

The cells were measured for GFP fluorescence using a BD Celesta cytometer. GFP was excited with a 488nm laser and emission captured at 513 $\pm$ 26nm. Forward and side scatter gating was initially used to capture cell-sized events and discard small debris. Events positive for GFP were determined by gating at the minimum level for which the negative control cell samples (cells not treated with lentivirus) showed <0.5% of events positive for GFP expression. The gated cells positive for GFP fluorescence were then assessed for the % of GFP-positive cells of the total cells. To calculate lentiviral functional titer, a transduced cell well showing a GFP % positive that was between 5% and 20% of cells was used to determine titer. The formula for virus titer calculation: $\text{titer} = \{(F \times C_n) / V\} \times DF$. F is the frequency of GFP-positive cells determined by flow cytometry; C_n = The total number of target cells infected. V = The volume of the inoculum, and DF = dilution factor. As shown in Fig. 4, CD8 binders 46, 1, 52, and 17 showed similar titer in human CD8 α and human CD8 $\alpha\beta$ overexpressing cell lines, indicating that those binders are specific for CD8 α . However, CD8 binder 97 had a significantly higher titer on human CD8 $\alpha\beta$ overexpressing cell line as compared to the cell line expressing only CD8 α , indicating that CD8 binder 97 is specific for CD8 β .

Binding assessment and affinity vs human CD8 $\alpha\beta$ heterodimer

Binders were generated as crude preparations of scFvs with a v5 tag and captured onto a Streptavidin BLI sensor that was functionalized with an anti-V5 tag antibody

for analysis via a ForteBio Octet HTX instrument. Sensors were exposed to 300nM of human CD8 $\alpha\beta$ followed by dissociation into assay buffer. A positive binding response was referenced by buffer subtraction with a negative control scFv and fit using a 1:1 Langmuir model to determine off-rates.

A subset of scFvs were produced as recombinant, purified scFv with a mIgG2a Fc tag. scFv-mIgG2a proteins were loaded onto anti-mouse Fc capture sensors for analysis via ForteBio Octet RED96 instrument. Sensors were exposed to a serial dilution of human CD8 alpha beta starting at 500nM, followed by dissociation into assay buffer. A positive binding response was referenced by buffer subtraction and fit using a 1:1 Langmuir model to determine on and off-rates and calculate a K_D .

Binding assessment vs human CD8 α

Binders were generated as crude preparations of scFvs with a v5 tag and captured onto a HC200M SPRi sensor that was functionalized with an anti-V5 tag antibody for analysis via a Cytiva LSA instrument. A serial dilution of human CD8 α starting at 5 μ M was injected over the sensor, followed by an injection of assay buffer for dissociation. A positive binding response was referenced by buffer subtraction with a negative control scFv and fit using a 1:1 Langmuir model to determine binding response.

Binding assessment vs human CD8 β

A subset of scFvs were produced as recombinant, purified scFv or VHH with a mIgG2a Fc tag. Human CD8 β protein was biotinylated and loaded onto a Streptavidin sensor for analysis via ForteBio Octet RED96 instrument. Sensors were exposed to a serial dilution of concentrations of scFv-mIgG2a or VHH-mIgG2a starting at 500nM, followed by dissociation into assay buffer. A positive binding response was referenced by buffer subtraction and fit using a 1:1 Langmuir model to determine on and off-rates and calculate a K_D .

Binding assessment vs cyno CD8 α homodimer

Binders were generated as crude preparations of scFvs with a v5 tag and captured onto a HC200M SPRi sensor that was functionalized with an anti-V5 tag antibody for analysis via a Cytiva LSA instrument. A serial dilution of cyno CD8 α -hFc starting at 1 μ M was injected over the sensor, followed by an injection of assay buffer for

dissociation. A positive binding response was referenced by buffer subtraction with a negative control scFv and fit using a 1:1 Langmuir model to determine off-rates. A subset of scFvs and VHH were produced as recombinant, purified scFv or VHH with a mIgG2a Fc tag. scFv-mIgG2a or VHH-mIgG2a proteins were loaded onto anti-mouse Fc capture sensors for analysis via ForteBio Octet RED96 instrument. Sensors were exposed to a serial dilution of concentrations of cyno CD8 α -hFc starting at 500nM, followed by dissociation into assay buffer. A positive binding response was referenced by buffer subtraction and fit using a 1:1 Langmuir model to determine on and off-rates and calculate a K_D .

The binding affinities of select CD8 binders are shown in **Table 16**.

Table 16. Binding summary of select CD8 scFvs.

CD8 Binder	Crude Titers Sup1T1 (TU/mL)	Highest Response (nm)	K_D (M)	k_{on} (1/Ms)	k_{off} (1/s)
46	4.90E+06	0.4906	9.52E-08	8.63E+04	8.21E-03
1	1.05E+06	0.1334	2.03E-08	3.08E+04	6.23E-04
5	1.83E+06	0.2201	4.01E-07	3.84E+04	1.54E-02
97	3.47E+06	0.4651	5.30E-09	5.14E+04	2.72E-04

Example 2: Generation and characterization of viral vectors pseudotyped with CD8 binders

Methods for generating Nipah G pseudotyped viruses

This example describes methods to generate and characterize functional titers of Nipah G/F pseudo-typed lentivirus functionalized with a CD8-specific scFv or VHH. Eliminating binding of Nipah G to EphrinB2/B3 and functionalizing Nipah G with a CD8-specific binder allows for targeting of CD8-expressing cells with the lentiviral vector. Nipah G (d34, E501A;W504A;Q530A;E533A) was modified by the addition of a glycine/serine linker (G4Sx3) followed by a CD8-specific binder (scFv or VHH) at the C-terminus. The Nipah G – linker -binder construct was codon optimized for expression in human cells and sub-cloned into an expression vector for lentivirus generation.

Crude lentiviral production was performed as follows: HEK-293LX cells were plated 24 hours in advance of transfection. On the day of transfection, HEK-293LX cells

were transfected with the lentiviral packaging plasmid (psPAX2), the lentiviral transfer plasmid encoding GFP (pSFFV-GFP), and the plasmids encoding for Nipah G protein retargeted for CD8 receptor targeting (NiV-G(CD8)) and Nipah F fusion protein (NiV-Fd22). After 24 hours, a media change was performed, and after another 24 hours, the lentivirus was harvested. To harvest the lentivirus, supernatant was removed from the HEK293LX cells and spun at 1000xg for 5 minutes. The supernatant was removed and immediately added to CD8-positive target cells or T cells, or frozen at -80 °C for later use.

Concentrated lentiviral production was performed as follows: LV-Max HEK293 cells were seeded 24 hours in advance of transfection. On the day of transfection, LV-Max HEK293 cells were transfected with the lentiviral packaging plasmid (psPAX2), the lentiviral transfer plasmid encoding GFP (pSFFV-GFP), and the plasmids encoding for Nipah G protein retargeted for CD8 receptor targeting (NiV-G(CD8)) and Nipah F fusion protein (NiV-Fd22). Two days later supernatant from the transfected cells was harvested, passed through a 0.45 µm filter, and concentrated by ultracentrifugation at 120,000 x g for 90 minutes. After ultracentrifugation the lentiviral pellet was resuspended in PBS, and either used immediately for transduction or aliquoted and frozen at -80°C for use later.

Transduction efficiency on human and nemestrina primary cells

Lentiviral production was performed as follows: LV-Max HEK293 cells were seeded 24 hours in advance of transfection. On the day of transfection, LV-Max HEK293 cells were transfected with the lentiviral packaging plasmid (psPAX2), the lentiviral transfer plasmid encoding GFP (pSFFV-GFP), and the plasmids encoding for Nipah G protein retargeted for CD8 receptor targeting (NiV-G(CD8)) and Nipah F fusion protein (NiV-Fd22). Two days later supernatant from the transfected cells was harvested, passed through a 0.45 µm filter, and concentrated by ultracentrifugation at 120,000 x g for 90 minutes. After ultracentrifugation the lentiviral pellet was resuspended in PBS, and either used immediately for transduction or aliquoted and frozen at -80°C for use later.

Primary isolated Human PanT cells or PBMCs were obtained from StemCell Technologies and activated with anti-human CD3/CD28 with IL-2 supplemented in the media. M. nemestrina PBMCs were obtained from BioIVT and activated using

anti-CD3/CD28 beads from Miltenyi Biotec with IL-2 supplemented in the media. 72 hours after activation, cells were harvested and plated in a 96-well plate. A serial dilution of lentivirus was performed and added to the primary PanT or PBMC cells. A spinfection was performed at 1000xg for 90 minutes at 25°C. The cells were then incubated at 37°C with 5% CO₂ for 72 hours. After 72 hours, cells were harvested via centrifugation at 1000xg for 5 minutes and stained with a live/dead exclusionary stain and an antibody panel to identify CD3 positive, CD8 positive and CD4 positive populations.

The cells were measured for GFP fluorescence using a BD Celesta cytometer. Forward and side scatter gating was initially used to capture cell-sized events and discard small debris. Forward scatter height and forward scatter width was used to capture single cells. Events negative for live/dead stain were selected. For PanT cells, cells were gated on either CD8 positive and GFP positive or CD4 negative and GFP positive to assess transduction in CD8-expressing cells. GFP expression in CD8 positive cells was calculated based upon the percent of cells expressing GFP from a CD8 positive population or CD4 negative population. Events positive for GFP were determined by gating at the minimum level for which the negative control cell samples (cells not treated with lentivirus) showed <0.5% of events positive for GFP expression. To calculate lentiviral functional titer, a transduced cell well showing a GFP % positive of CD8 positive cells that was between 5% and 20% of cells was used to determine titer. The formula for virus titer calculation: $\text{titer} = \{(F \times C_n) / V\} \times DF$. F is the frequency of GFP-positive cells determined by flow cytometry; C_n = The total number of target cells infected. V = The volume of the inoculum, and DF = dilution factor.

For PBMC cells, cells were measured for GFP fluorescence using a BD Fortessa cytometer. Cells were gated to be CD3 positive, and then gated for either CD8 positive and GFP positive or CD4 negative and GFP positive to assess transduction in CD8-expressing cells. GFP expression in CD8 positive cells was calculated based upon the percent of cells expressing GFP from a CD8 positive population or CD4 negative population. Events positive for GFP were determined by gating at the minimum level for which the negative control cell samples (cells not treated with lentivirus) showed <0.5% of events positive for GFP expression. To calculate lentiviral functional titer, a transduced cell well showing a GFP positive % of CD8

positive cells that was between 5% and 20% of cells was used to determine titer. The formula for virus titer calculation: $\text{titer} = \{(F \times C_n) / V\} \times DF$. F is the frequency of GFP-positive cells determined by flow cytometry; C_n = The total number of target cells infected. V = The volume of the inoculum, and DF = dilution factor (see Figs. 5-8).

In vivo specificity

To demonstrate specific in vivo transduction, GFP-expressing lentivirus with CD8 binders 97 and 46 were produced as above and administered to mice. T cells in human PBMCs were activated for 3 days via CD3/CD28 treatment and 1E7 cells were intraperitoneally injected into NOD-scid-IL2 γ^{null} mice. One day later 1E7 TU of CD8 fusosome (binder 97 or 46) or vehicle control were intraperitoneally injected. 7 days later peritoneal cells were harvested by peritoneal lavage and analyzed by flow cytometry for on-target GFP expression and off-target expression. Approximately 35-55% of the CD8⁺ cells in the mice were GFP positive following treatment with the CD8-targeted lentiviral vectors (**Fig. 9A**), whereas less than 0.2% of (**Fig. 9B**) were GFP positive, indicating CD8-specific transduction in the mice.

Tumor killing

To evaluate tumor killing in vitro, human PBMC were activated with anti-CD3/CD28 reagent and cells were transduced with CD8-targeted fusosomes (VHH binder 97 or scFv binder 46) expressing CD19CAR or GFP. RFP⁺ Nalm6 cells were added to the cultures on day 3, and elimination of Nalm6 cells was evaluated by flow cytometry 18 hours after CD19CAR expression was detected specifically in CD8⁺ cells with both CD8 fusogens 4 days after transduction (**Fig. 10A**). Fusogen-generated CD19CAR T cells (solid symbols) mediated potent and fusosome dose-dependent killing of CD19-expressing Nalm6 cells. In contrast, fusogen generated GFP⁺ T cells (open symbols) did not exhibit target cell killing (**Fig. 10B**).

Example 3: In vivo delivery of a CD8 targeted fusogens in Nalm6 tumor models as a function of T cell activation state

This Example describes the assessment of the transduction efficiency of CD8 retargeted Nipah fusogens and VSV-G.

Briefly, sixty-two (62) female NSG mice were injected with 1E6 Nalm6-Luc leukemia B cells via intravenous (IV) injection, followed three days later by an IV injection of 2E6 human peripheral blood mononuclear cells (hPBMC), with or without prior T cell activation with CD3/CD28 complexes. A day after hPBMC injection, CD8 binder 97 Nipah fusogen pseudotyped lentiviral vector (LV) expressing a CD19 CAR were injected at a range of integrating units (IU), 2E6-5E7, into separate groups of animals. Nalm6 tumor progression was tracked via bioluminescent imaging (BLI) using the Lago X imaging system weekly throughout the duration of the study. The CD19CAR contained an anti-scFv directed against CD19 and an intracellular signaling domain containing intracellular components of 4-1BB and CD3-zeta. Peripheral blood analysis was performed on half of the mice from each group to assess circulating T and B cell frequencies, circulating CAR-T cell frequencies, and cytokine levels throughout the duration of the study. The study was concluded 28 days post-hPBMC injection, or earlier based on individual animal health. Animals were sacrificed and cells from peripheral blood, spleen, and bone marrow tissues were harvested and analyzed by flow cytometry for CD19CAR expressing cells and cytokine analysis.

As shown in **Fig. 11A**, CD8-CD19CAR LV and activated hPBMC treatment resulted in robust control of Nalm6 tumor growth over time. As shown in **Fig. 11B**, high dose CD8-CD19CAR LV and non-activated hPBMC treatment resulted in slightly delayed yet robust control of Nalm6 tumor growth. **Fig. 11C** shows the percent of on-target CD19CAR expressing cells (CD8⁺CD19CAR⁺) in total recovered live lymphocytes as indicated in top right quadrant of the FACs plots in both PBMC control (top plots) and CD8 fusosome-treated animals (bottom plots). There was no statistical difference in the frequency of CAR-T cells in the bone marrow of animals that received 5 E7 IU of CD8-CD19CAR LV, either with or without hPBMC activation. The results indicate that CD19-specific CAR T cells could be detected in CD8⁺ T cells up to 28 days post-treatment in the peripheral blood, spleen and bone marrow.

Example 4: A 5 and 7-week Single Dose Pharmacokinetic and Pharmacodynamic Study of CD8-SFFV-CD20CAR by Intravenous Infusion in Juvenile Female Nemestrina macaques

This Example describes a lentiviral vector pseudotyped with an anti-CD8 binding protein (binder 46) targeting CD8+ T cells to deliver a CD20 CAR transgene (CD8-SFFV-CD20CAR). The CD20CAR contained an anti-scFv directed against CD20 and an intracellular signaling domain containing intracellular components of 4-1BB and CD3-zeta. The objective of this study was to characterize the ability of CD8-SFFV-CD20CAR to transduce T cells and deplete normal, healthy CD20+ B cells, the biodistribution of viral integration, and tolerability of intravenous administration. Eight juvenile female nemestrina macaques were administered CD8-SFFV-CD20CAR at a single maximum feasible dose of 7.69E8 IU/kg (n=6) or saline control at 10 ml/kg (n=2) intravenously over 1 hour. Animals were evaluated at baseline (Day -35, -28 and -21) for frequency of B and T-cells together with an assessment of hematological and clinical chemistry parameters. On-study animals were monitored daily for clinical observations, weekly for changes in body weight, temperature, neurological battery, and hematology and clinical chemistry. CSF samples were collected pre-study, on Day 7 and at termination (from days 35-52). All animals underwent routine blood sampling and flow cytometry immune-phenotyping for changes in B and T-cell frequencies on Day 3, 5, 7, 10, 14, 17, 21, 28, and 35. At termination, animals underwent a full necropsy, blood, CSF, and tissues were harvested for: flow cytometry of lymphoid tissues, cytokine analyses by Luminex, transgene expression by PCR, vector copy number (VCN) by ddPCR, insertion site distribution (ISD) by deep sequencing, clinical pathology (hematology and clinical chemistry), tissue immunohistochemistry, and anatomic histopathology. Interim data to Day 35 demonstrated that administration CD8-SFFV-CD20CAR at a single maximum feasible dose of 7.69E8 IU/kg was well-tolerated in all animals. There were no compound-related changes in clinical observations including neurological signs, body temperature nor clinical chemistry values across all sampling times. There were transient minimal reductions of platelets and neutrophils on Day 7-10 that returned to baseline by Day 14. There were transient, minimal decreases in hematocrit and associated increase in reticulocyte counts in two

animals, which may be attributed to repeated blood sampling. As shown in **Fig. 12**, flow cytometric analyses demonstrated a significant decrease in CD20+ B cells in 4 of 6 treated animals beginning on Day 7 in peripheral blood compared to intra-animal pre-dosing, that was sustained through Day 35. Normalized B cell counts through day 28/35 are presented in **Fig. 13**. These data are consistent with the anticipated pharmacological activity of the anti-CD20CAR.

Preliminary VCN measurements using ddPCR were performed on samples from 2 control and 4 treated animals from Day -35 and Day 14 and 35 in peripheral blood mononuclear cells (PBMC's), and at termination in spleen and bone marrow. At Day -35, and 35 post-injection, VCN in PBMC's was observed though the values were below limit of quantitation (BLQ) in all animals. By comparison in the spleen VCN was detected in treated animals whereas control animals were BLQ. At Day 35, 0.04 to 1.3% of splenocytes (ie. 67 to 1,970 cells) contained at least one inserted copy in CD8-SFFV-CD20CAR treated animals examined. CD20 CAR mRNA was also detected in the spleen of the treated monkeys but not in the control animals.

Additional VCN measurements of transgene integration in the genome was measured via droplet digital polymerase chain reaction (ddPCR) using the WPRE (Woodchuck hepatitis virus Post-transcriptional Response Element) amplicon.

Transgene was quantifiable or detectable in PBMCs of all (6 of 6) fusosome-dosed animals collected on study days 3 through 9/10 (**Table 17**), and detectable on days 14 through days 27/28 in PBMC samples in Animals 1, 4, and 6.

Table 17: Transgene integration measured by vector copy number (VCN) for PBMC samples collected in-life.

Animal	Day -28	Day 3	Day 5	Day 7	Days 9/10	Day 14	Days 16-18	Day 21	Days 27/28	Day 35
Control 1	U	U	U	U	U	U	U	U	U	U
Control 2	U	U	U	U	U	No sample	U	U	U	U
1	U	0.00342	0.00273	D	0.00220	U	D	D	D	U
2	U	D	D	0.00207	0.00263	U	U	U	U	U
3	U	D	D	D	0.00283	U	U	U	U	U
4	U	0.00331	D	0.00326	0.00230	D	D	U	U	U
5	U	0.00254	0.00179	0.00289	0.00131	U	U	U	U	U
6	U	0.00820	0.00718	No Sample	No Sample	D	D	D	U	U

Vector copy number (VCN) was only reported when WPRE copies were in range of quantitation (50 to 120,000 copies per reaction). VCN was calculated as WPRE copies x2/TERT and reported as VCN (integrations per diploid genome). If WPRE VCN copies were below LLOQ (Lower limit of

quantification of <50 copies per reaction) but above LLOD (Lower limit of detection of <10 copies per reaction) WPRE VCN was reported as detectable. If WPRE copies were below WPRE VCN was reported undetectable. D, detectable; U, undetectable.

These data demonstrate on-target activity of CD8-SFFV-CD20CAR in immune competent animals that was well tolerated and was correlated with presence of vector in cells in the spleen, even in the absence of administration of T-cell activating treatment.

Table 5. VH Sequences

CD8 Binder	VH Sequence	SEQ ID NO:
1	QVQLVQSGAEVKKPGASVKVSCKASGGTFSSYAISWVRQAPGQGLEWMGIIDPSDGNNTYQAQ NFKQGRVTMTRDTSTSTVYMELSSLRSEDAVYYCAKERAAGYMDVWGQGTITVTVSS	405
2	QVQLVQSGAEVKKPGASVKVSCKASGGTFNTYAINWVRQAPGQGLEWMGRIDPSSGGTKYQAQ NFKQGRVTMTRDTSTSTVYMELSSLRSEDAVYYCAKEHAAGTYMDVWGKGTITVTVSS	406
3	QVQLVQSGAEVKKPGASVKVSCKASGGTFSSYAINWVRQAPGQGLEWMGIIDPSGGNTYQAQ NFKQGRVTMTRDTSTSTVYMELSSLRSEDAVYYCAKERAAGYMDVWGQGTITVTVSS	407
4	QVQLVQSGAEVKKPGASVKVSCKASGYTFTGYMHVVRQAPGQGLEWMGHINPNNGDTNY AQNFQGRVTMTRDTSTSTVYMELSSLRSEDAVYYCAKEGYMDVWGQGTITVTVSS	651
5	QVQLVQSGAEVKKPGASVKVSCKASGYTFTDYIQWVRQAPGQGLEWMGWINPNSGGTSYAQ QKFQGRVTMTRDTSTSTVYMELSSLRSEDAVYYCAKEGDDYMDAWGQGTMTVTVSS	408
6	QVQLVQSGAEVKKPGASVKVSCKASGYTFTYDIHWVRQAPGQGLEWMGVINPNDGSTRYAQ NFKQGRVTMTRDTSTSTVYMELSSLRSEDAVYYCARERGGMPDYWGQGTITVTVSS	409
7	QVQLVQSGAEVKKPGASVKVSCKASGYTFTSYAMNWVRQAPGQGLEWMGRINPNSGGTNYA QKFQGRVTMTRDTSTSTVYMELSSLRSEDAVYYCARGHGIPKYWGQGTITVTVSS	410
8	QVQLVQSGAEVKKPGASVKVSCKASGYTFTSYIHWVRQAPGQGLEWMGWMNPNSGNTGY AQKFQGRVTMTRDTSTSTVYMELSSLRSEDAVYYCARVRSGSPQHWGQGTITVTVSS	411
9	QVQLVQSGAEVKKPGASVKVSCKASGHTFSRHYIHWVRQAPGQGLEWMGWMNPNSGNTGY AQKFQGRVTMTRDTSTSTVYMELSSLRSEDAVYYCARGGPWIVDAFDIWGQGTMTVTVSS	412
10	QVQLVQSGAEVKKPGASVKVSCKASGYTFTSYGISWVRQAPGQGLEWMGWISAHNGVTQYA QKFQGRVTMTRDTSTSTVYMELSSLRSEDAVYYCARGIAVAGTDYWGQGTITVTVSS	652
11	QVQLVQSGAEVKKPGASVKVSCKASGGTFSNTDINWVRQAPGQGLEWMGIINPSGGTSYAQ KFQGRVTMTRDTSTSTVYMELSSLRSEDAVYYCAREATWGPYYMDVWGKGTITVTVSS	413
12	QVQLVQSGAEVKKPGASVKVSCKASGYTFTRSYVHWVRQAPGQGLEWMGWISPYNGNTKYA QKFQGRVTMTRDTSTSTVYMELSSLRSEDAVYYCVRNKDGLQHWGQGTITVTVSS	414
13	QVQLVQSGAEVKKPGASVKVSCKASGDTFTGYMHVVRQAPGQGLEWMGIINPNSGDTKYA HQFQGRVTMTRDTSTSTVYMELSSLRSEDAVYYCAKDAKRVGYMDVWGKGTITVTVSS	415
14	QVQLVQSGAEVKKPGASVKVSCKASGYTFTRYMHVVRQAPGQGLEWMGRINPNSGGTNYA QKFQGRVTMTRDTSTSTVYMELSSLRSEDAVYYCARLVGGSPDYWGQGTITVTVSS	416
15	QVQLVQSGAEVKKPGASVKVSCKASGYTFTNYDINWVRQAPGQGLEWMGRINPNSGGTNYAE NFKQGRVTMTRDTSTSTVYMELSSLRSEDAVYYCARGAMVDYWGQGTITVTVSS	417
16	QVQLVQSGAEVKKPGASVKVSCKASGGTFSNTDINWVRQAPGQGLEWMGIINPSDGDTKYAQ EFQGRVTMTRDTSTSTVYMELSSLRSEDAVYYCARGNYVGSYYYMDVWGQGTITVTVSS	418
17	QVQLVQSGAEVKKPGASVKVSCKASGYTFTNYLHWVRQAPGQGLEWMGWINPNSGDTKYA QKFQGRVTMTRDTSTSTVYMELSSLRSEDAVYYCARDSRGDWYFDLWGRGTITVTVSS	419
18	QVQLVQSGAEVKKPGASVKVSCKASGYGFTSYIHWVRQAPGQGLEWMGVIDPSSGGTSYAQ KFQGRVTMTRDTSTSTVYMELSSLRSEDAVYYCTRHGGRLADYWGQGTITVTVSS	420
19	QVQLVQSGAEVKKPGASVKVSCKASGYTFTSRDISWVRQAPGQGLEWMGWIDPKSGDITYA KFQGRVTMTRDTSTSTVYMELSSLRSEDAVYYCARLKESSILDAFDIWGQGTMTVTVSS	421
20	QVQLVQSGAEVKKPGASVKVSCKASGYTFTSYDINWVRQAPGQGLEWMGMINPGAGSSTYAQ KFQGRVTMTRDTSTSTVYMELSSLRSEDAVYYCARERFGTGYMDVWGQGTMTVTVSS	422
21	EVQLLESQGGGLVQPGGSLRLSCAASGFTFSNDSMNWVRQAPGKLEWVSLISGDDGTTYADS VKGRFTISRDNKNTLYLQMNSLRAEDAVYYCARVIGEMVDDAFDLWGQGTITVTVSS	423
22	QVQLVQSGAEVKKPGASVKVSCKASGYTFTGYMHVVRQAPGQGLEWMGSINPNSGDTGYA QKFQGRVTMTRDTSTSTVYMELSSLRSEDAVYYCARERLFGTYMDVWGKGTITVTVSS	424
23	QVQLVQSGAEVKKPGSSVKVSCKASGYTFTYDINWVRQAPGQGLEWMGRIPIFGTANYAQKF QGRVTITADESTSTAYMELSSLRSEDAVYYCARADGELTDYWGQGTITVTVSS	425
24	EVQLLESQGGGLVQPGGSLRLSCAASGFTFSYTMWVRQAPGKLEWVSAIGTGGGIYADSVK GRFTISRDNKNTLYLQMNSLRAEDAVYYCARHHLPAHYMDVWGKGTITVTVSS	426
25	QVQLVQSGAEVKKPGASVKVSCKASGGTFSRYDINWVRQAPGQGLEWMGRINPNSGDTNYA QKFQGRVTMTRDTSTSTVYMELSSLRSEDAVYYCARDVPAGRYMDVWGKGTITVTVSS	427

26	QVQLVQSGAEVKKPGASVKVSCKASGNTFTSYMHVVRQAPGQGLEWMGMINPSDGSTRYA QKFQGRVTMTRDTSTSTVYMELSSLRSEDAVYYCAKDRGVGRYYYYMDVWGKTTTVSS	428
27	QVQLVQSGAEVKKPGASVKVSCKASGGTFSRYAVSWVRQAPGQGLEWMGIINPSDGSTTYAQ KFQGRVTMTRDTSTSTVYMELSSLRSEDAVYYCAKDSRYGRYYYYMDVWGKTTTVSS	429
28	QVQLVQSGAEVKKPGSSVKVSCKASGGTFSNYAISWVRQAPGQGLEWMGIINPNGGSPSYAQK FQGRVTITADESTSTAYMELSSLRSEDAVYYCAKEIVVGPYYYYMDVWGKTTTVSS	430
29	QVQLVQSGAEVKKPGSSVKVSCKASGGTFTRYAISWVRQAPGQGLEWMGRINPNSGDTNYAQ KFQGRVTITADESTSTAYMELSSLRSEDAVYYCARGMVRGPYYYYMDVWGKTTTVSS	431
30	QVQLVQSGAEVKKPGSSVKVSCKASGGTFSYIAISWVRQAPGQGLEWMGIINPSGGSTSYAQTF QGRVTITADESTSTAYMELSSLRSEDAVYYCAREGVTGPYYYYMDVWGQGTTLTVSS	432
31	QVQLVQSGAEVKKPGASVKVSCKASGGTFSRFIDINWVRQAPGQGLEWMGIINPSDGSTDY AQ NFQGRVTMTRDTSTSTVYMELSSLRSEDAVYYCARDAAAGTRYYYYYGMDVWGQGTTLTVSS	433
32	QVQLVQSGAEVKKPGASVKVSCKASGGTFSHAISWVRQAPGQGLEWMGIINPSGGSTSYAQK FQGRVTMTRDTSTSTVYMELSSLRSEDAVYYCARELYSSTYYYYMDVWGKTTTVSS	434
33	QVQLVQSGAEVKKPGASVKVSCKASGGTFSYIAISWVRQAPGQGLEWMGRINPNTGGTNHAQ KFQGRVTMTRDTSTSTVYMELSSLRSEDAVYYCARALYSGPYYYYYMDVWGKTTTVSS	435
34	EVQLLESGLLVQPGGSLRLSCAASGFTFSNDSMNWVRQAPGKGLEWVSAISGSGSTYYADS VKGRFTISRDN SKNTLYLQMNSLRAEDAVYYCAKEHAAGTYYYYYMDVWGKTTTVSS	436
35	QVQLVQSGAEVKKPGASVKVSCKASGGTFSYGINWVRQAPGQGLEWMGWISGYN GDTDYA RKLQGRVTMTRDTSTSTVYMELSSLRSEDAVYYCARDSLVGRYYYYMDVWGKTTTVSS	437
36	QVQLVQSGAEVKKPGASVKVSCKASGYFTDYDIYWVRQAPGQGLEWLGWISADNGNTNYEQ KVQGRVTMTRDTSTSTVYMELSSLRSEDAVYYCARRSELDYWGQGTTLTVSS	438
37	QVQLVQSGAEVKKPGASVKVSCKASGYFTSYMHVVRQAPGQGLEWMGWISPN SGATHYA QKFQGRVTMTRDTSTSTVYMELSSLRSEDAVYYCARGDDNDYWGQGTTLTVSS	439
38	QVQLVQSGAEVKKPGASVKVSCKASGYFTSYDINWVRQAPGQGLEWMGWINPNSGNTGYA KKFQGRVTMTRDTSTSTVYMELSSLRSEDAVYYCARGEEDYWGQGTTLTVSS	440
39	QVQLVQSGAEVKKPGASVKVSCKASGYFTSYPMNWVRQAPGQGLEWMGIINPSGGSTTYAQ KFQGRVTMTRDTSTSTVYMELSSLRSEDAVYYCARGRRVPDYWGQGTTLTVSS	441
40	QVQLVQSGAEVKKPGASVKVSCKASGYFTDYIHWVRQAPGQGLEWMGWINPKSGATNYA QKFQGRVTMTRDTSTSTVYMELSSLRSEDAVYYCARGKVTTDYWGQGTTLTVSS	442
41	EVQLLESGLLVQPGGSLRLSCAASGFTFSSFMNWVRQAPGKGLEWVSRISESGDSSFYADSV KGRFTISRDN SKNTLYLQMNSLRAEDAVYYCASGRELIEYWGQGTTLTVSS	443
42	EVQLLESGLLVQPGGSLRLSCAASGFTFDDYAMHWVRQAPGKGLEWVSAIGTGGGTYADS VKGRFTISRDN SKNTLYLQMNSLRAEDAVYYCARVYDFPDVWGQGTTLTVSS	444
43	QVQLVQSGAEVKKPGASVKVSCKASGYFTDSYMHVVRQAPGQGLEWMGWMNPSNGDTG YARKFQGRVTMTRDTSTSTVYMELSSLRSEDAVYYCARSTYSHIDYWGQGTTLTVSS	445
44	QVQLVQSGAEVKKPGASVKVSCKASGYFTSYMHVVRQAPGQGLEWMGTISPDSGSTTYAQ RFQGRVTMTRDTSTSTVYMELSSLRSEDAVYYCAREDSGFDYWGQGTTLTVSS	446
45	QVQLVQSGAEVKKPGASVKVSCKASGYTFMNYIHWVRQAPGQGLEWMGIINPSGGSTTYAQ KFQGRVTMTRDTSTSTVYMELSSLRSEDAVYYCARDQGGGFDYWGQGTTLTVSS	447
46	QVQLVQSGAEVKKPGASVKVSCKASGYFTSYMHVVRQAPGQGLEWMGGFDPEDGETIYA QKFQGRVTMTRDTSTSTVYMELSSLRSEDAVYYCARDQGWGMDVWGQGTTLTVSS	448
47	QVQLVQSGAEVKKPGASVKVSCKASGYFTSYIHWVRQAPGQGLEWMGRINPKSGRTTYAQ NFQGRVTMTRDTSTSTVYMELSSLRSEDAVYYCARLTEGIPDYWGQGTTLTVSS	449
48	QVQLVQSGAEVKKPGASVKVSCKASGYTLNDYIHWVRQAPGQGLEWMGVINPGGGSTTYAQ TFQGRVTMTRDTSTSTVYMELSSLRSEDAVYYCARDRYGPFYWGQGTTLTVSS	450
49	QVQLVQSGAEVKKPGASVKVSCKASGYFTSYDINWVRQAPGQGLEWMGLMNPKTGDTNYA EKFQGRVTMTRDTSTSTVYMELSSLRSEDAVYYCTRLVAGGAPDYWGQGTTLTVSS	451
50	QVQLVQSGAEVKKPGASVKVSCKASGYFTGYMHVVRQAPGQGLEWMGIIDPSDGYTSYAQ KFQGRVTMTRDTSTSTVYMELSSLRSEDAVYYCARDGFTGDIAIWGQGTTLTVSS	452
51	QVQLVQSGAEVKKPGSSVKVSCKASGYFTGYMHVVRQAPGQGLEWMGWINPNSGGTNY AQKFQGRVTITADESTSTAYMELSSLRSEDAVYYCARVDDSSPDYWGQGTTLTVSS	453
52	QVQLVQSGAEVKKPGSSVKVSCKASGYFTGYLHWVRQAPGQGLEWMGGIMPISGTTIYAQK FQGRVTITADESTSTAYMELSSLRSEDAVYYCTTGPDGTEVDYWGQGTTLTVSS	454

52	QVQLVQSGAEVKKPGASVKVSCKASGYFTTNHYMHVVRQAPGQGLEWMGWMNPNSGNTG YAQKFQGRVTMTRDTSTSTVYMESSLRSEDVAVYYCASSESGSDLDYWGQGLTVTVSS	455
54	QVQLVQSGAEVKKPGASVKVSCKASGYFTNYYIHVVRQAPGQGLEWMGWMSPSGDTGYA QKFQGRVTMTRDTSTSTVYMESSLRSEDVAVYYCAREVEIEGYMDVWVGQGTTVTVSS	456
55	QVQLVQSGAEVKKPGASVKVSCKASGYFTTSYYMHVVRQAPGQGLEWMGWINPNSGDTSYA QKFQGRVTMTRDTSTSTVYMESSLRSEDVAVYYCAKDLDDWYMDVWVGKGTTVTVSS	457
56	QVQLVQSGAEVKKPGASVKVSCKASGYFTTSYYMHVVRQAPGQGLEWMGIIDPSGDITSYAQK FQGRVTMTRDTSTSTVYMESSLRSEDVAVYYCTDSTTWDAFDIWGQGTMTVTVSS	458
57	QVQLVQSGAEVKKPGASVKVSCKASGYFTDYMHVVRQAPGQGLEWMGWINPNSGGTNY AQKFQGRVTMTRDTSTSTVYMESSLRSEDVAVYYCARVLVSGSPDYWGQGLTVTVSS	459
58	QVQLVQSGAEVKKPGASVKVSCKASGYFTTENEMHVVRQAPGQGLEWMGIETSGGSTDYAQ KFQGRVTMTRDTSTSTVYMESSLRSEDVAVYYCAREAAAGLDFQHWGQGLTVTVSS	460
59	QVQLVQSGAEVKKPGASVKVSCKASGYTFASYDMHVVRQAPGQGLEWMGIINPNSGGTNYA QKLQGRVTMTRDTSTSTVYMESSLRSEDVAVYYCARANSWDAMVIDYWGQGLTVTVSS	461
60	EVQLLESGGGLVQPGGSLRLSCAASGFTFSNSDMHVVRQAPGKGLEWVSISGSGVTYYADS VKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCAREHSSSWYTFDYWGQGLTVTVSS	462
61	QVQLVQSGAEVKKPGASVKVSCKASGYFTTAYMHVVRQAPGQGLEWLWINPNSGGTDYA QKFQGRVTMTRDTSTSTVYMESSLRSEDVAVYYCARDDDSSGYLDYWGQGLTVTVSS	463
62	QVQLVQSGAEVKKPGASVKVSCKASGYFTNYYIHVVRQAPGQGLEWMGMINPSGGSTTYAQ KFQGRVTMTRDTSTSTVYMESSLRSEDVAVYYCARASGDYMDLIDYWGQGLTVTVSS	464
63	QVQLVQSGAEVKKPGASVKVSCKASGYFTDYMHVVRQAPGQGLEWLWINPDSGGTNYE QKFQGRVTMTRDTSTSTVYMESSLRSEDVAVYYCALVGSGLYAPTHWGQGLTVTVSS	465
64	QVQLVQSGAEVKKPGSSVKVSCKASGYPTDYMHVVRQAPGQGLEWMGWMNPNSGNTG YAQKFQGRVTITADESTSTAYMESSLRSEDVAVYYCARVRGDGYNLGDYWGQGLTVTVSS	466
65	QVQLVQSGAEVKKPGASVKVSCKASGYTFSDYMHVVRQAPGQGLEWMGWINPNSGGTNS AQKFQGRVTMTRDTSTSTVYMESSLRSEDVAVYYCARDVDTAMGAGDYWGQGLTVTVSS	467
66	QVQLVQSGAEVKKPGASVKVSCKASGYFTDYIHVVRQAPGQGLEWIGIINPSGGSAYAQKF QGRVTMTRDTSTSTVYMESSLRSEDVAVYYCARVARWGYGDYPDYWGQGLTVTVSS	468
67	QVQLVQSGAEVKKPGASVKVSCKASGDTFTTHDINWVRQAPGQGLEWMGIISPSDGSTSYAQK LQGRVTMTRDTSTSTVYMESSLRSEDVAVYYCARDNRNGDYGYMDVWVGQGTTVTVSS	469
68	QVQLVQSGAEVKKPGASVKVSCKASGDTFTNYYIHVVRQAPGQGLEWMGWINPISGGTHYA QKFQGRVTMTRDTSTSTVYMESSLRSEDVAVYYCAREGLSSWYVLDYWGQGLTVTVSS	470
69	QVQLVQSGAEVKKPGASVKVSCKASGYFTTSYDINWVRQAPGQGLEWMGWISADNGDTSFA QKFQGRVTMTRDTSTSTVYMESSLRSEDVAVYYCARDGSHYGYGYMDVWVGQGTTVTVSS	471
70	QVQLVQSGAEVKKPGSSVKVSCKASGYFTTSYDINWVRQAPGQGLEWMGGISPIYGTPAYAQK FQGRVTITADESTSTAYMESSLRSEDVAVYYCASPGPEGYYGYMDVWVGQGTTVTVSS	472
71	QVQLVQSGAEVKKPGASVKVSCKASGYFTDNHYMHVVRQAPGQGLEWMGWMNPNSGNTG YAQKFQGRVTMTRDTSTSTVYMESSLRSEDVAVYYCASYHWDYGDYRFDYWGQGLTVTVSS	473
72	QVQLVQSGAEVKKPGASVKVSCKASGYFTTSYYIHVVRQAPGQGLEWMGWMNPNSGNTGY AQKFQGRVTMTRDTSTSTVYMESSLRSEDVAVYYCARVEIDYGDSPPDYWGQGLTVTVSS	474
73	QVQLVQSGAEVKKPGASVKVSCKASGGTSSSYAISWVRQAPGQGLEWMGIINPSDGDTSYAQK FQGRVTMTRDTSTSTVYMESSLRSEDVAVYYCARGAEWELRYAFDIWGQGTMTVTVSS	475
74	QVQLVQSGAEVKKPGASVKVSCKASGYFTTYDISWVRQAPGQGLEWMGTINPSGGTITYAQK FQGRVTMTRDTSTSTVYMESSLRSEDVAVYYCARETYGLYYGYMDVWVGKGTTVTVSS	476
75	QVQLVQSGAEVKKPGASVKVSCKASGYFTTSYDINWVRQAPGQGLEWMGWMNPKSGNTGY AQKFQGRVTMTRDTSTSTVYMESSLRSEDVAVYYCARAPSLRGYSYGPDYWGQGLTVTVSS	477
76	QVQLVQSGAEVKKPGASVKVSCKASGGTFTSYDINWVRQAPGQGLEWMGIINPSGGSTSYAQK FQGRVTMTRDTSTSTVYMESSLRSEDVAVYYCAKDRQERYYYGYMDVWVGKGLTVTVSS	478
77	QVQLVQSGAEVKKPGASVKVSCKASGYFTTSYDINWVRQAPGQGLEWMGIINPSDGDSTDYAQK FQGRVTMTRDTSTSTVYMESSLRSEDVAVYYCAKDRSYGDYGYMDVWVGQGTTVTVSS	479
78	QVQLVQSGAEVKKPGASVKVSCKASGGTFTSYDINWVRQAPGQGLEWMGIINPGGNGARHTQ KFQGRVTMTRDTSTSTVYMESSLRSEDVAVYYCAREVFSENYYGYMDVWVGKGTTVTVSS	480
79	QVQLVQSGAEVKKPGASVKVSCKASGYFTTSYYMHVVRQAPGQGLEWMGIINPSDGGSTTYAQ KFQGRVTMTRDTSTSTVYMESSLRSEDVAVYYCAREWDYTHYYGYMDVWVGQGTTVTVSS	481

80	QVQLVQSGAEVKKPGASVKVSCKASGNTFTSHWIHWVRQAPGQGLEWMGGFDPEDGETVYA QNFQGRVTMTRDTSTSTVYMELSSLRSEDTAVYYCARGDSSGGYYQYFDYWGQGLTVTVSS	482
81	QVQLVQSGAEVKKPGSSVKVSCKASGYTFTSYDINWVRQAPGQGLEWMGGITPVFGIANYAQK FQGRVTITADESTSTAYMELSSLRSEDTAVYYCARGSWDSSSWIPEYWGQGLTVTVSS	483
82	QVQLVQSGAEVKKPGASVKVSCKASGFTFSDYDIVWVRQAPGQGLEWMGIINPRGGSTNYAQ KFQGRVTMTRDTSTSTVYMELSSLRSEDTAVYYCASLVWGGAYYYMDVWVGQGTTVTVSS	484
83	QVQLVQSGAEVKKPGASVKVSCKASGYTFTSYGISWVRQAPGQGLEWMGWMNPNNGDTDY AQKFQGRVTMTRDTSTSTVYMELSSLRSEDTAVYYCTTPVFSGSYWYFDPWGQGLTVTVSS	485
84	QVQLVQSGAEVKKPGASVKVSCKASGYTFTSYDINWVRQAPGQGLEWMGIINPSGGSTSYAQK FQGRVTMTRDTSTSTVYMELSSLRSEDTAVYYCTTDQAVAGPYYYGMDVWVGQGLTVTVSS	486
85	QVQLVQSGAEVKKPGASVKVSCKASGGTFSSYAIWVRQAPGQGLEWMGLINPGSGNTNYAQ KFQGRVTMTRDTSTSTVYMELSSLRSEDTAVYYCARDRWLAGPYYYGMDVWVGQGTTVTVSS	487
86	QVQLVQSGAEVKKPGSSVKVSCKASGYMFTGHDMHWVRQAPGQGLEWMGGIIPFGTPNYA QKFQGRVTITADESTSTAYMELSSLRSEDTAVYYCARVMGPPVDDYYGMDVWVGQGTTVTVSS	488
87	QVQLVQSGAEVKKPGASVKVSCKASGYIFSNYDMHWVRQAPGQGLEWMGIINPSDGSSTYAQ KFQGRVTMTRDTSTSTVYMELSSLRSEDTAVYYCARDLGPFGSYYYMDVWVGKGTTVTVSS	489
88	EVQLVESGGGLVKPGGSLRLSCAASGFTFSSYAMTWVRQAPGKGLEWVSTINGDGDSTDYADS VKGRFTISRDDSKNTLYLQMNSLKTEDTAVYYCAREGVVPPYYYYMDVWVGKGTTVTVSS	490
89	QVQLVQSGAEVKKPGASVKVSCKASGYTFTTYMHWVRQAPGQGLEWMGQIDPNSGDTIYP QKFQGRVTMTRDTSTSTVYMELSSLRSEDTAVYYCARSSGWSRYYYYMDVWVGKGTTVTVSS	491
90	QVQLVQSGAEVKKPGASVKVSCKASGSTFTNYQIHWVRQAPGQGLEWMGIINPSGGSTSYAQK FQGRVTMTRDTSTSTVYMELSSLRSEDTAVYYCARDNGMTTGYYYMDVWVGKGTTVTVSS	492
91	QVQLVQSGAEVKKPGASVKVSCKASGYTFTSYDIVWVRQAPGQGLEWMGIINPSGGSTSYAQK FQGRVTMTRDTSTSTVYMELSSLRSEDTAVYYCARDAMVTGYYYGMDVWVGQGTTVTVSS	493
92	QVQLVQSGAEVKKPGASVKVSCKASGYTFTSYDINWVRQAPGQGLEWMGIVNPSDGNNTYAQ KFQGRVTMTRDTSTSTVYMELSSLRSEDTAVYYCARDRGYDGRGYYGMDVWVGQGTTVTVSS	494
93	QVQLVQSGAEVKKPGASVKVSCKASGGTSSYDINWVRQAPGQGLEWMGWINTYNGNTYYA QKLQGRVTMTRDTSTSTVYMELSSLRSEDTAVYYCATSPKATADYYYYMDVWVGKGTTVTVSS	495
94	QVQLVQSGAEVKKPGASVKVSCKASGYTFTSYDINWVRQAPGQGLEWMGIINPSDGITDYAQR FQGRVTMTRDTSTSTVYMELSSLRSEDTAVYYCTTSTVTPSYYYGMDVWVGQGTTVTVSS	496
95	QVQLVQSGAEVKKPGASVKVSCKASGYTFTSHAIHWVRQAPGQGLEWMGIINPRDGDTVYAQ KFQGRVTMTRDTSTSTVYMELSSLRSEDTAVYYCAREPVAGTGYYYGMDVWVGQGLTVTVSS	497
96	QVQLVQSGAEVKKPGASVKVSCKASGGTFNSYGINWVRQAPGQGLEWMGWMNPNNSGNTGY AQKFQGRVTMTRDTSTSTVYMELSSLRSEDTAVYYCARDNLGAFWSDYYYYGMDVWVGQGTTV TVSS	498
97		

Table 6. VL Sequences

CD8 Binder	VL Sequence	SEQ ID NO:
1	DIQMTQSPSSLSASVGDRVITICRASQSISSYLNWYQQKPKGAPKLLIYAASSLQSGVPSRFSGSG SGTDFTLTISSLQPEDFATYYCQQSYSTPLTFGGGKTKVEIKR	499
2	DIQMTQSPSSLSASVGDRVITICQASQDISYLNWYQQKPKGAPKLLIYAASSLQSGVPSRFSGS GSGTDFTLTISSLQPEDFATYYCQQSYSNLVFSGGQTKVEIKR	500
3	DIQMTQSPSSLSASVGDRVITICRASQSISSYLNWYQQKPKGAPKLLIYAASSLQSGVPSRFSGSG SGTDFTLTISSLQPEDFATYYCQQSYSTPLTFGGGKTKVEIKR	501
4	DIVMTQSPLSLPVTGPGEASISCRSSQSLLHNSGYNLYDWYLQKPGQSPQLLIYLSNRASGVDPDR FSGSGSGTDFTLKISRVEAEDVGVYYCMQALQTPFTFGPGTKVDIKR	502
5	DIVMTQSPLSLPVTGPGEASISCRSSQSLLHNSGYNLYDWYLQKPGQSPQLLIYLSNRASGVDPDR FSGSGSGTDFTLKISRVEAEDVGVYYCMQGLQTPHTFGGQTKVEIKR	503
6	DIQMTQSPSSLSASVGDRVITICRASQISRNINWYQQKPKGAPKLLIYKASNLKGGVPSRFSGS GSGTDFTLTISSLQPEDFATYYCQQTYSAPLFGGQTKLEIKR	504

7	DIVMTQSPSLPVPTEGEPASISCRSSQSLHNSGYNLDWYLQKPGQSPQLLIYLSNRASGVDPDR FSGSGSGTDFTLKISRVEADVGVYYCMTLQTLPTFGQGKTKEIKR	505
8	EIVMTQSPATLSVSPGERATLSCRASQSVASDLAWYQQKPGQAPRLLIYGASTRATGIPARFSGS GSGTEFTLTISLQSEDFAVYYCQYQDPSGSGGQGTKEIKR	506
9	DIQMTQSPSSLSASVGDRVITITCQASQDIGNYLNWYQQKPGKAPKLLIYAASLTQRGVPSRFSGS GSGTDFLTISLQPEDFATYYCQQAANSFPPTFGGGTKVEIKR	507
10	EIVMTQSPATLSVSPGERATLSCRASQSISTHLAWYQQKPGQAPRLLIYGASTRATGIPARFSGSG SGTEFTLTISLQSEDFAVYYCQYQNSRTFGGQGTKEIKR	508
11	DIQMTQSPSSLSASVGDRVITITCRASQTISNYLNWYQQKPGKAPKLLIYAASLTQSGVPSRFSGSG SGTDFLTISLQPEDFATYYCQSYSTPPTFGQGKTKEIKR	509
12	DIQMTQSPSSLSASVGDRVITITCRASQGIKNDLWYQQKPGKAPKLLIYDASTLQSGVPSRFSGS GSGTDFLTISLQPEDFATYYCQSYSSPYTFGGQGTKEIKR	510
13	DIQMTQSPSSLSASVGDRVITITCRASQSISNYLNWYQQKPGKAPKLLIYAASLTQSGVPSRFSGSG SGTDFLTISLQPEDFATYYCQSYSTPYTFGGQGTKEIKR	511
14	DIVMTQSPSLPVPTEGEPASISCRSSQSLHNSGYNLDWYLQKPGQSPQLLIYLSNRASGVDPDR FSGSGSGTDFTLKISRVEADVGVYYCMTLQGAHWPPPTFGQGKTKEIKR	512
15	DIQMTQSPSSLSASVGDRVITITCRASQGISDLSAWYQQKPGKAPKLLIYGASSLRSGVPSRFSGSG SGTDFLTISLQPEDFATYYCQSYRTPYTFGGQGTKEIKR	513
16	DIQMTQSPSSLSASVGDRVITITCRASQSISNYLNWYQQKPGKAPKLLIYAASLTQSGVPSRFSGSG SGTDFLTISLQPEDFATYYCQSFSTQWTFGGQGTKEIKR	514
17	DIQMTQSPSSLSASVGDRVITITCQASQDIHNYLNWYQQKPGKAPKLLIYDASLETGVPSRFSGS GSGTDFLTISLQPEDFATYYCQQAANSFPPTFGPGTKVDIKR	515
18	DIQMTQSPSSLSASVGDRVITITCQASQDISNYLNWYQQKPGKAPKLLIYAASLTQSGVPSRFSGS GSGTDFLTISLQPEDFATYYCQQRNWPYTFGGQGTKEIKR	516
19	DIQMTQSPSSLSASVGDRVITITCRASQISDWLAWYQQKPGKAPKLLIYAASLTQGVPSRFSGS GSGTDFLTISLQPEDFATYYCQQAISFPTFGQGTKEIKR	517
20	DIQMTQSPSSLSASVGDRVITITCQASQDISNYLNWYQQKPGKAPKLLIYAASLTQSGVPSRFSGS GSGTDFLTISLQPEDFATYYCQSYSPPTFGPGTKVDIKR	518
21	DIQMTQSPSSLSASVGDRVITITCRASQSISTWLAWYQQKPGKAPKLLIYAASLTQSGVPSRFSGS GSGTDFLTISLQPEDFATYYCQQAISFPLTFGGGTKEIKR	519
22	DIQMTQSPSSLSASVGDRVITITCRASQSISNYLNWYQQKPGKAPKLLIYAASLTQSGVPSRFSGSG SGTDFLTISLQPEDFATYYCQSYTFPTFGGQGTKEIKR	520
23	DIVMTQSPSLPVPTEGEPASISCRSSQSLHNSGYNLDWYLQKPGQSPQLLIYDASHLETGVDPDR FSGSGSGTDFTLKISRVEADVGVYYCQYQSYSPPTFGQGKTKEIKR	521
24	DIQMTQSPSSLSASVGDRVITITCQASQDISNYLNWYQQKPGKAPKLLIYAASLTQSGVPSRFSGS GSGTDFLTISLQPEDFATYYCQQSYSAPLTFPGGTKEIKR	522
25	DIQMTQSPSSLSASVGDRVITITCQASQDISNYLNWYQQKPGKAPKLLIYAASLTQSGVPSRFSGS GSGTDFLTISLQPEDFATYYCQQSFSTFYTFGGQGTKEIKR	523
26	DIQMTQSPSSLSASVGDRVITITCQASQDISNYLNWYQQKPGKAPKLLIYAASLTQSGVPSRFSGS GSGTDFLTISLQPEDFATYYCQQSYSPPTFGPGTKVDIKR	524
27	DIQMTQSPSSLSASVGDRVITITCRASQSINRFLNWYQQKPGKAPKLLIYAASLTQNGVPSRFSGS GSGTDFLTISLQPEDFATYYCQSYSTPYTFGGQGTKEIKR	525
28	DIQMTQSPSSLSASVGDRVITITCRASQSISSYLNWYQQKPGKAPKLLIYAASLTQSGVPSRFSGSG SGTDFLTISLQPEDFATYYCQSYSTPLTFGGGTKEIKR	499
29	DIQMTQSPSSLSASVGDRVITITCQASQDISNYLNWYQQKPGKAPKLLIYAASLTQSGVPSRFSGS GSGTDFLTISLQPEDFATYYCQQSYSTPITFGGQGTKEIKR	526
30	DIQMTQSPSSLSASVGDRVITITCRASQSVSTYLNWYQQKPGKAPKLLIYAASLTQSGVPSRFSGS GSGTDFLTISLQPEDFATYYCQQSYTIPSTFGGQGTKEIKR	527
31	DIQMTQSPSSLSASVGDRVITITCQASQDIKYNWYQQKPGKAPKLLIYAASLTQSGVPSRFSGS GSGTDFLTISLQPEDFATYYCQQSYSAAPTFGGGTKVEIKR	528
32	DIQMTQSPSSLSASVGDRVITITCQASQGITNYLNWYQQKPGKAPKLLIYGASSLTQSGVPSRFSGS GSGTDFLTISLQPEDFATYYCQQSYSTPWTFPGGTKEIKR	529
33	DIQMTQSPSSLSASVGDRVITITCRASQSISSYLNWYQQKPGKAPKLLIYAASLTQSGVPSRFSGSG SGTDFLTISLQPEDFATYYCQQSYSTPLTFGGGTKEIKR	499

34	DIQMTQSPSSLSASVGDRVITITCQASQDIHNYLNWYQQKPGKAPKLLIYAASLTQSGVPSRFSGS GSGTDFTLTISSLQPEDFATYYCQSYTTPLTFGGGTKVEIKR	530
35	DIQMTQSPSSLSASVGDRVITITCQASQDISNYLNWYQQKPGKAPKLLIYSAFSLQSGVPSRFSGS GSGTDFTLTISSLQPEDFATYYCQSYSAPITFGGQTRLEIKR	531
36	DIQMTQSPSSLSASVGDRVITITCRASQSISSYLNWYQQKPGKAPKLLIYASNLQSGVPSRFSGS SGTDFTLTISSLQPEDFATYYCQQRSNWPPVTFGGGTKVEIKR	532
37	DIQMTQSPSSLSASVGDRVITITCQANQDISNLEWYQQKPGKAPKLLIYDASSLESVPSRFSGS SGTDFTLTISSLQPEDFATYYCQSYSIPITFGGQTRLEIKR	533
38	DIQMTQSPSSLSASVGDRVITITCRASQGISNNLNWYQQKPGKAPKLLIYEASTLESVPSRFSGS SGTDFTLTISSLQPEDFATYYCQSYSTPLTFGGGTKVEIKR	534
39	DIVMTQSPLSLPVTGPASISCRSSQSLHNSGNYLDWYLQKPGQSPQLLIYGASTLETGPDR FSGSGSGTDFTLKISRVEAEDVGYYCMQGLQPPGTFGGGTKVEIKR	535
40	DIQMTQSPSSLSASVGDRVITITCRASQISRSVLWYQQKPGKAPKLLIYAASLTQGVPSRFSGS SGTDFTLTISSLQPEDFATYYCQSYNHFRFTFGGQTKVDIKR	536
41	DIQMTQSPSSLSASVGDRVITITCQASQDISNYLNWYQQKPGKAPKLLIYDASLETGPVPSRFSGS GSGTDFTLTISSLQPEDFATYYCQQRSDSTPLTFGGGTKVEIKR	537
42	DIQMTQSPSSLSASVGDRVITITCQASHDISLNWYQQKPGKAPKLLIYGASTLQSGVPSRFSGS GSGTDFTLTISSLQPEDFATYYCQQLNSYPRTFGGGTKVEIKR	538
43	DIQMTQSPSSLSASVGDRVITITCRASQDIGAYLWYQQKPGKAPKLLIYAASLTQSGVPSRFSGS GSGTDFTLTISSLQPEDFATYYCQSYSIPYTFGGGTKLEIKR	539
44	DIQMTQSPSSLSASVGDRVITITCRASQSISSYLWYQQKPGKAPKLLIYAASLTQSGVPSRFSGS SGTDFTLTISSLQPEDFATYYCQSYSTPYTFGGGTKLEIKR	540
45	DIQMTQSPSSLSASVGDRVITITCRASQGISYLAWYQQKPGKAPKLLIYGASNLETGPVPSRFSGS SGTDFTLTISSLQPEDFATYYCQSYSTPYTFGGGTKLEIKR	541
46	DIQMTQSPSSLSASVGDRVITITCRASQSISSYLNWYQQKPGKAPKLLIYAASLTQSGVPSRFSGS SGTDFTLTISSLQPEDFATYYCQQTSTPYTFGGGTKLEIKR	542
47	DIQMTQSPSSLSASVGDRVITITCRASQIGNITWLAWYQQKPGKAPKLLIYAASLTQSGVPSRFSGS GSGTDFTLTISSLQPEDFATYYCQSYSTPQTTFGGGTKVDIKR	543
48	DIQMTQSPSSLSASVGDRVITITCRASQTISSYLNWYQQKPGKAPKLLIYAASLTQSGVPSRFSGS SGTDFTLTISSLQPEDFATYYCQSYRTPYTFGGGTKLEIKR	544
49	DIVMTQSPLSLPVTGPASISCRSSQSLHNSGNYLDWYLQKPGQSPQLLIYMGSNRASGVDP RFSGSGSGTDFTLKISRVEAEDVGYYCMQGTHWPTFGGQTRLEIKR	545
50	DIQMTQSPSSLSASVGDRVITITCRASQINNNYLNWYQQKPGKAPKLLIYGASSLTQSGVPSRFSGS GSGTDFTLTISSLQPEDFATYYCQQTFSLPYTFGGGTKVEIKR	546
51	DIQMTQSPSSLSASVGDRVITITCRASQTISSYLNWYQQKPGKAPKLLIYDASLETGPVPSRFSGS SGTDFTLTISSLQPEDFATYYCQSYSTPYTFGGGTKLEIKR	547
52	DIQMTQSPSSLSASVGDRVITITCRASRGIGNDLAWYQQKPGKAPKLLIYDASTLETGPVPSRFSGS GSGTDFTLTISSLQPEDFATYYCQQGYNMPLTFGGGTKVEIKR	548
52	DIQMTQSPSSLSASVGDRVITITCRASQTISSYLNWYQQKPGKAPKLLIYGASNLTGPVPSRFSGS GSGTDFTLTISSLQPEDFATYYCQQTYSAPLTFGGGTKVEIKR	549
54	DIQMTQSPSSLSASVGDRVITITCRASQFIGSWLAWYQQKPGKAPKLLIYAASLTQSGVPSRFSGS GSGTDFTLTISSLQPEDFATYYCQSYSPWTFGGGTKVEIKR	550
55	DIQMTQSPSSLSASVGDRVITITCRASQSISSWMAWYQQKPGKAPKLLIYDASLETGPVPSRFSGS GSGTDFTLTISSLQPEDFATYYCQQTSTPYTFGGGTKVEIKR	551
56	DIQMTQSPSSLSASVGDRVITITCRASQGISNNLNWYQQKPGKAPKLLIYDASLETGPVPSRFSGS GSGTDFTLTISSLQPEDFATYYCQQSYSSPWTFGGGTKVEIKR	552
57	DIVMTQSPDLSLAVSLGERATINCKSSQSVLYSSNNKNYLAWYQQKPGQPPKLLIYWASTRESGVP DRFSGSGSGTDFTLTISSLQAEDVAVYYCQYASAPRTFGGQTKLEIKR	553
58	DIQMTQSPSSLSASVGDRVITITCRASQSISSYLNWYQQKPGKAPKLLIYKTSSLESVPSRFSGS GTDFTLTISSLQPEDFATYYCQQSFTIPYTFGGGTKVEIKR	554
59	DIQMTQSPSSLSASVGDRVITITCRVSSQSISSYLNWYQQKPGKAPKLLIYGASSLTQSGVPSRFSGS SGTDFTLTISSLQPEDFATYYCQSYSTPLTFGGGTKVEIKR	555
60	DIQMTQSPSSLSASVGDRVITITCRASQISDWLAWYQQKPGKAPKLLIYDASLETGPVPSRFSGS GSGTDFTLTISSLQPEDFATYYCQSYSTPLTFGGGTKVEIKR	556

61	DIQMTQSPSSLSASVGDRVITITCRASQGISNYLAWYQQKPGKAPKLLIYSASNLSQSGVPSRFSGS GSGDFTLTISSLQPEDFATYYCQQTYPPTFGPGTKVDIKR	557
62	DIQMTQSPSSLSASVGDRVITITCRASQISIRNYLTWYQQKPGKAPKLLIYSASNLSQSGVPSRFSGS SGTDFTLTISSLQPEDFATYYCQQSYSTPLTFGQGTKVEIKR	558
63	DIQMTQSPSSLSASVGDRVITITCRASQIRNYLTWYQQKPGKAPKLLIYAASLTQSGVPSRFSGS SGTDFTLTISSLQPEDFATYYCQQLTTPFTFGPGTKVDIKR	559
64	DIQMTQSPSSLSASVGDRVITITCRASQDIRKFLNWWYQQKPGKAPKLLIYAASLTQSGVPSRFSGS GSGDFTLTISSLQPEDFATYYCQQLNGYPGTGFGQGTKLEIKR	560
65	DIQMTQSPSSLSASVGDRVITITCRASQISSYLNWYQQKPGKAPKLLIYTASNLSQSGVPSRFSGS SGTDFTLTISSLQPEDFATYYCQQSYSTPLTFGGGKTKVEIKR	561
66	DIQMTQSPSSLSASVGDRVITITCRASQDISNYLWYQQKPGKAPKLLIYDASNLSQSGVPSRFSGS GSGDFTLTISSLQPEDFATYYCQQTYPPTFGPGTKVDIKR	562
67	DIQMTQSPSSLSASVGDRVITITCRASQNVRSWLAWYQQKPGKAPKLLIYAASLTQSGVPSRFSGS GSGDFTLTISSLQPEDFATYYCQQSYNTPTTFGQGTKLEIKR	563
68	DIQMTQSPSSLSASVGDRVITITCRASQIGNDLGWYQQKPGKAPKLLIYAASLTQSGVPSRFSGS GSGDFTLTISSLQPEDFATYYCQQSYAPPTFGQGTKVEIKR	564
69	DIQMTQSPSSLSASVGDRVITITCRASQISNWLAWYQQKPGKAPKLLIYGASNLETGVPFSRFSGS GSGDFTLTISSLQPEDFATYYCQQSYSTPPTFGQGTKLEIKR	565
70	DIVMTQSPSLPVPTEGEPASISCRSSQSLHSGNYLDWYQKPGQSPQLLIYLSNRASGVDPDR FSGSGSGDFTLTKISRVEAEDVGVYCMQGLQTPFTFGQGTKVEIKR	566
71	DIQMTQSPSSLSASVGDRVITITCRASQISSYLNWYQQKPGKAPKLLIYLASSLSQSGVPSRFSGS GTDFTLTISSLQPEDFATYYCQQSDSIPVTFGQGTKVEIKR	567
72	DIQMTQSPSSLSASVGDRVITITCRASQDISNYLWYQQKPGKAPKLLIYSTSSLSQSGVPSRFSGS SGTDFTLTISSLQPEDFATYYCQQSYSTPYNFGQGTKLEIKR	568
73	DIQMTQSPSSLSASVGDRVITITCRASEIGSWLAWYQQKPGKAPKLLIYAASLTQSGVPSRFSGS SGTDFTLTISSLQPEDFATYYCQQSYSTPTTFGQGTKLEIKR	569
74	DIQMTQSPSSLSASVGDRVITITCRASQISNWLAWYQQKPGKAPKLLIYAASLTQSGVPSRFSGS SGTDFTLTISSLQPEDFATYYCQQSYSTPLTFGQGTKVEIKR	570
75	EIVMTQSPATLSVSPGERATLSCRASQSVTSNYLAWYQQKPGQAPRLLIYGASTRATGIPARFSGS GSGTEFTLTISSLQSEDFAVYYCQHYGSSPAFGQGTKLEIKR	571
76	DIQMTQSPSSLSASVGDRVITITCRASQISSYLNWYQQKPGKAPKLLIYAASLTQSGVPSRFSGS SGTDFTLTISSLQPEDFATYYCQQSYSTPLTFGGGKTKVEIKR	499
77	DIQMTQSPSSLSASVGDRVITITCRASQISSYLNWYQQKPGKAPKLLIYAASLTQSGVPSRFSGS SGTDFTLTISSLQPEDFATYYCQQSYSTPPTFGPGTKVDIKR	572
78	DIQMTQSPSSLSASVGDRVITITCRASQDIGNYLNWYQQKPGKAPKLLIYAASLTQSGVPSRFSGS GSGDFTLTISSLQPEDFATYYCQQAAYTYPPTFGQGTKLEIKR	573
79	DIQMTQSPSSLSASVGDRVITITCRASQDISNYLWYQQKPGKAPKLLIYGASSLSQSGVPSRFSGS GSGDFTLTISSLQPEDFATYYCQQSYTTPNTFGPGTKVDIKR	574
80	DIQMTQSPSSLSASVGDRVITITCRASQGISNYLAWYQQKPGKAPKLLIYAASLTQSGVPSRFSGS SGTDFTLTISSLQPEDFATYYCQQSYSTPTTFGQGTKVEIKR	575
81	DIQMTQSPSSLSASVGDRVITITCRASQISNGLSWYQQKPGKAPKLLIYDASNLETGVPFSRFSGS GSGDFTLTISSLQPEDFATYYCQQSYSTPFTFGPGTKVDIKR	576
82	DIQMTQSPSSLSASVGDRVITITCRASQIRNYLTWYQQKPGKAPKLLIYGASSLSQSGVPSRFSGS GSGDFTLTISSLQPEDFATYYCQQSYSTPLTFGGGKTKVEIKR	577
83	DIQMTQSPSSLSASVGDRVITITCRASQDISNYLWYQQKPGKAPKLLIYKASSLESQSGVPSRFSGS SGTDFTLTISSLQPEDFATYYCQQSYSMPLTFGPGTKVDIKR	578
84	DIQMTQSPSSLSASVGDRVITITCRASQDISNYLWYQQKPGKAPKLLIYAASLTQSGVPSRFSGS GSGDFTLTISSLQPEDFATYYCQQSYTTPWTFGQGTKLEIKR	579
85	DIQMTQSPSSLSASVGDRVITITCRASQDISNYLWYQQKPGKAPKLLIYAASLTQSGVPSRFSGS GSGDFTLTISSLQPEDFATYYCQQSYSSPLTFGGGKTKVEIKR	580
86	DIQMTQSPSSLSASVGDRVITITCRASQDISNYLWYQQKPGKAPKLLIYKASSLESQSGVPSRFSGS SGTDFTLTISSLQPEDFATYYCQQSYSDPLTFGQGTKVEIKR	581
87	DIQMTQSPSSLSASVGDRVITITCRASQDISNYLWYQQKPGKAPKLLIYGASTLSQSGVPSRFSGS GSGDFTLTISSLQPEDFATYYCQQSYAPITFGQGTKLEIKR	582

88	DIQMTQSPSSLSASVGDRVITITCRASQISNYLNWYQQKPGKAPKLLIYAASNLQSGVPSRFSGS GSGTDFLTISLQPEDFATYYCQSYTTPFTFGPGTKVDIKR	583
89	DIQMTQSPSSLSASVGDRVITITCRASQIGNYLNWYQQKPGKAPKLLIYAASNLQSGVPSRFSGS GSGTDFLTISLQPEDFATYYCQSYSTPPWTFGQGTKEIKR	584
90	DIQMTQSPSSLSASVGDRVITITCRASQISNYLNWYQQKPGKAPKLLIYAASNLQSGVPSRFSGS GSGTDFLTISLQPEDFATYYCQSYQTPLTFGGGKVEIKR	585
91	DIQMTQSPSSLSASVGDRVITITCRASQISNYLNWYQQKPGKAPKLLIYAASNLQSGVPSRFSGS GSGTDFLTISLQPEDFATYYCQSYTTPFTFGQGTKEIKR	586
92	DIQMTQSPSSLSASVGDRVITITCRASQISNYLNWYQQKPGKAPKLLIYAASNLQSGVPSRFSGS GSGTDFLTISLQPEDFATYYCQSYSTPQTFGQGTKEIKR	587
93	DIQMTQSPSSLSASVGDRVITITCRASQIRNDLNWYQQKPGKAPKLLIYAASNLQSGVPSRFSGS GSGTDFLTISLQPEDFATYYCQANSFPTFGQGTKEIKR	588
94	DIQMTQSPSSLSASVGDRVITITCRASQINTWLAWYQQKPGKAPKLLIYAASNLQSGVPSRFSGS GSGTDFLTISLQPEDFATYYCQSYSTPYTFGQGTKEIKR	589
95	DIQMTQSPSSLSASVGDRVITITCRASQISNYLNWYQQKPGKAPKLLIYAASNLQSGVPSRFSGS GSGTDFLTISLQPEDFATYYCQSYTVPPTFGQGTKEIKR	590
96	DIQMTQSPSSLSASVGDRVITITCRASQIDIRYFLNWYQQKPGKAPKLLIYAASNLQSGVPSRFSGS GSGTDFLTISLQPEDFATYYCQDDSFPLTFGGGKVEIKR	591

Table 7

SEQ ID NO:	SEQUENCE	ANNOTATION
592	MVVILDKRCY CNLLILILMI SECSVGILHY EKLSKIGLVK GVTRKYKIKS NPLTKDIVIK MIPNVSNMSQ CTGSVMENYK TRLNGILTPI KGALEIYKNN THDLVGDVRL AGVIMAGVAI GIATAAQITA GVALYEAMKN ADNINKLKSS IESTNEAVVK LQETAECTVY VLTALQDYIN TNLVPTIDKI SCKQTELSLD LALSKYLSDL LFVFGPNLQD PVSNSMTIQA ISQAFGGNYE TLLRTLGYAT EDFDDLLESD SITGQIIYVD LSSYYIIVRV YFPILTEIQQ AYIQELLPVS FNNDNSEWIS IVPNFILVRN TLISNIEIGF CLITKRSVIC NQDYATPMTN NMRECLTGST EKCPRELVS SHVPRFALS GVLFANCISV TCQCQTGRA ISQSGETLL MIDNTTCPTA VLGNIISLG KYLGSVNYS EGIAIGPPVF TDKVDISSQI SSMNQLQQS KDYIKEAQL LDTVNPSLIS MLSMIILYVL SIASLCIGLI TFISFIIVEK KRNTYSRLED RVRPTSSGD LYYIGT	Nipah virus NiV-F with signal sequence (aa 1-546) Uniprot Q9IH63
593	ILHY EKLSKIGLVK GVTRKYKIKS NPLTKDIVIK MIPNVSNMSQ CTGSVMENYK TRNLGILTPI KGALEIYKNN THDLVGDVRL AGVIMAGVAI GIATAAQITA GVALYEAMKN ADNINKLKSS IESTNEAVVK LQETAECTVY VLTALQDYIN TNLVPTIDKI SCKQTELSLD LALSKYLSDL LFVFGPNLQD PVSNSMTIQA ISQAFGGNYE TLLRTLGYAT EDFDDLLESD SITGQIIYVD LSSYYIIVRV YFPILTEIQQ AYIQELLPVS FNNDNSEWIS IVPNFILVRN TLISNIEIGF CLITKRSVIC NQDYATPMTN NMRECLTGST EKCPRELVS SHVPRFALS GVLFANCISV TCQCQTGRA ISQSGETLL MIDNTTCPTA VLGNIISLG KYLGSVNYS EGAIGPPVF TDKVDISSQI SSMNQLQQS KDYIKEAQL LDTVNPSLIS MLSMIILYVL SIASLCIGLI TFISFIIVEK KRNTYSRLED RVRPTSSGD LYYIGT	Nipah virus NiV-F F0 (aa 27-546)

594	ILHYEKLKSKIGLVKGVTRKYKIKSNPLTKDIVIKMIPNVSNMSQ CTGSMENYKTRNLGILTPIKGALEIYKNNTHDLVGDVR	Nipah virus NiV-F F2 (aa 27-109)
595	LAGVIMAGVAIGIATAAQITAGVALYEAMKNADNINKLKSSIES TNEAVVKLQETAECTVYVLTALQDYINTNLVPTIDKISCKQTEL SLDLALSKYLSDDL FVFGPNLQDPVSNMTIQAISQAFGGNYET LLRTLGYATEDFDDLLESDSITGQIIYVDLSSYYIIVRVYFPIL TEIQQAYIQELLPVSFNNNDSEWISIVPNFILVRNTLISNIEIG FCLITKRSVICNQDYATPMTNNMRECLTGSTEKCPRELVVSSHV PRFALSNGVLFANCISVTCQCQTTGRAISQSSEGTLLMIDNTTC PTAVLGNVVIISLGKYLGSVNYNSEGIAIGPPVFTDKVDISSQIS SMNQSLQQSKDYIKEAQRLLDTVNPSLISMLSMIILYVLSIASL CIGLITFISFIIVEKKRNTYSRLEDRRVRPTSSGDLYYIGT	Nipah virus NiV F F1 (aa 110-546)
596	ILHY EKLKSKIGLVK GVTRKYKIKS NPLTKDIVIK MIPNVSNMSQ CTGSMENYK TRNLGILTPI KGALEIYKNN THDLVGDVRL AGVIMAGVAI GIATAAQITA GVALYEAMKN ADNINKLKSS IESTNEAVVK LQETAECTVY VLTALQDYIN TNLVPTIDKI SCKQTELSLD LALSKYLSDL LFVFGPNLQD PVSNSMTIQA ISQAFGGNYE TLLRTLGYAT EDFDDLLES SITGQIIYVD LSSYYIIVRV YFPILTEIQQ AYIQELLPV FNNNDSEWIS IVPNFILVRN TLISNIEIGF CLITKRSVIC NQDYATPMTN NMRECLTGST EKCPRELVVS SHVPRFALS N GVLFANCISV TCQCQTTGRA ISQSSEGTLL MIDNTTCPTA VLGNVVIISLG KYLGSVNYNS EGIAIGPPVF TDKVDISSQI SSMNQSLQQS KDYIKEAQRL LDTVNPSLIS MLSMIILYVL SIASLCIGLI TFISFIIVEK KRNTGT	Nipah virus NiV-F F0 truncation (aa 525- 544)
597	LAGVIMAGVAIGIATAAQITAGVALYEAMKNADNINKLKSSIES TNEAVVKLQETAECTVYVLTALQDYINTNLVPTIDKISCKQTEL SLDLALSKYLSDDL FVFGPNLQDPVSNMTIQAISQAFGGNYET LLRTLGYATEDFDDLLESDSITGQIIYVDLSSYYIIVRVYFPIL TEIQQAYIQELLPVSFNNNDSEWISIVPNFILVRNTLISNIEIG FCLITKRSVICNQDYATPMTNNMRECLTGSTEKCPRELVVSSHV PRFALSNGVLFANCISVTCQCQTTGRAISQSSEGTLLMIDNTTC PTAVLGNVVIISLGKYLGSVNYNSEGIAIGPPVFTDKVDISSQIS SMNQSLQQSKDYIKEAQRLLDTVNPSLISMLSMIILYVLSIASL CIGLITFISFIIVEKKRNTGT	Nipah virus NiV F F1 (aa 110-546) truncation (aa 525- 544)
598	ILHY EKLKSKIGLVK GVTRKYKIKS NPLTKDIVIK MIPNVSNMSQ CTGSMENYK TRNLGILTPI KGALEIYKNQ THDLVGDVRL AGVIMAGVAI GIATAAQITA GVALYEAMKN ADNINKLKSS IESTNEAVVK LQETAECTVY VLTALQDYIN TNLVPTIDKI SCKQTELSLD LALSKYLSDL LFVFGPNLQD PVSNSMTIQA ISQAFGGNYE TLLRTLGYAT EDFDDLLES SITGQIIYVD LSSYYIIVRV YFPILTEIQQ AYIQELLPV FNNNDSEWIS IVPNFILVRN TLISNIEIGF CLITKRSVIC NQDYATPMTN NMRECLTGST EKCPRELVVS SHVPRFALS N GVLFANCISV TCQCQTTGRA ISQSSEGTLL MIDNTTCPTA VLGNVVIISLG KYLGSVNYNS EGIAIGPPVF TDKVDISSQI SSMNQSLQQS KDYIKEAQRL LDTVNPSLIS MLSMIILYVL SIASLCIGLI TFISFIIVEK KRNTGT	Nipah virus NiV-F F0 truncation (aa 525- 544) AND mutation on N-linked glycosylation site
599	MVVIDDKRCY CNLLILILMI SECSVGILHY EKLSKIGLVK GVTRKYKIKS NPLTKDIVIK MIPNVSNMSQ CTGSMENYK TRNLGILTPI KGALEIYKNN THDLVGDVRL AGVIMAGVAI GIATAAQITA GVALYEAMKN ADNINKLKSS IESTNEAVVK LQETAECTVY VLTALQDYIN TNLVPTIDKI SCKQTELSLD LALSKYLSDL LFVFGPNLQD PVSNSMTIQA ISQAFGGNYE	Truncated NiV fusion glycoprotein (FcDelta22) at cytoplasmic tail (with signal sequence)

	TLLRTLGYAT EDFDDLLESD SITGQIIYVD LSSYYIIVRV YFPILTEIQQ AYIQELLPVS FNNDNSEWIS IVPNFILVRN TLISNIEIGF CLITKRSVIC NQDYATPMTN NMRECLTGST EKCPRELVVS SHVPRFALSN GVLFANCISV TCQCQTTGRA ISQSGEQTLL MIDNTTCPTA VLG NVIISLG KYLGSVNYNS EGIAIGPPVF TDKVDISSQI SSMNQSLQQS KDYIKEAQRL LDTVNPSLIS MLSMIILYVL SIASLCIGLI TFISFIIVEK KRNT	
1092	ILHY EKLSKIGLVK GVTRKYKIKS NPLTKDIVIK MIPNVSNSMQ CTGSVMENYK TRNGILTPI KGALEIYKNN THDLVGDVRL AGVIMAGVAI GIATAAQITA GVALYEAMKN ADNINKLKSS IESTNEAVVK LQETAECTVY VLTALQDYIN TNLVPTIDKI SCKQTELSLD LALSKYLSDL LFVFGPNLQD PVSNSMTIQA ISQAFGGNYE TLLRTLGYAT EDFDDLLESD SITGQIIYVD LSSYYIIVRV YFPILTEIQQ AYIQELLPVS FNNDNSEWIS IVPNFILVRN TLISNIEIGF CLITKRSVIC NQDYATPMTN NMRECLTGST EKCPRELVVS SHVPRFALSN GVLFANCISV TCQCQTTGRA ISQSGEQTLL MIDNTTCPTA VLGNVIISLG KYLGSVNYNS EGIAIGPPVF TDKVDISSQI SSMNQSLQQS KDYIKEAQRL LDTVNPSLIS MLSMIILYVL SIASLCIGLI TFISFIIVEK KRNT	Truncated NiV fusion glycoprotein (FcDelta22) F0
1093	LAGVIMAGVAI GIATAAQITA GVALYEAMKN ADNINKLKSS IESTNEAVVK LQETAECTVY VLTALQDYIN TNLVPTIDKI SCKQTELSLD LALSKYLSDL LFVFGPNLQD PVSNSMTIQA ISQAFGGNYE TLLRTLGYAT EDFDDLLESD SITGQIIYVD LSSYYIIVRV YFPILTEIQQ AYIQELLPVS FNNDNSEWIS IVPNFILVRN TLISNIEIGF CLITKRSVIC NQDYATPMTN NMRECLTGST EKCPRELVVS SHVPRFALSN GVLFANCISV TCQCQTTGRA ISQSGEQTLL MIDNTTCPTA VLG NVIISLG KYLGSVNYNS EGIAIGPPVF TDKVDISSQI SSMNQSLQQS KDYIKEAQRL LDTVNPSLIS MLSMIILYVL SIASLCIGLI TFISFIIVEK KRNT	Truncated NiV fusion glycoprotein (FcDelta22) F1
600	MGPAENKKVR FENTTSKDGK IPSKVIKSY GTMDIKKINE GLLDSKILSA FNTVIALLG IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLDRG DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCVST VGDPILNSTY WSGSLMTRL AVKPKSNGGG YNQHQALALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCRISM GIRPNSHYIL RSGLLKYNLS DGENPKVVF EISDQRLSIG SPSKIYDSLQ QPVFYQASFS WDTMIKFGDV LTVNPLVNW RNNTVISRPG QSQCPRNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC	NiVG protein attachment glycoprotein (602 aa)
601	MGKVR FENTTSKDGK IPSKVIKSY GTMDIKKINE GLLDSKILSA FNTVIALLG IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS	NiVG protein attachment glycoprotein Truncated Δ5

	CSRGVSKQRI IGVGEVLDRG DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCVAVST VGDPILNSTY WSGSLMMTRL AVKPKSNGGG YNQHQALALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCRLSM GIRPNSHYIL RSGLLKYNLS DGENPKVFI EISDQRLSIG SPSKIYDSLQ QPVFYQASFS WDTMIKFGDV LTVNPLVVNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC	
602	MGNNTSDKGG IPSKVIKSY GTMDIKKINE GLLDSKILSA FNTVIALGGS IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKIHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLDRG DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCVAVST VGDPILNSTY WSGSLMMTRL AVKPKSNGGG YNQHQALALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCRLSM GIRPNSHYIL RSGLLKYNLS DGENPKVFI EISDQRLSIG SPSKIYDSLQ QPVFYQASFS WDTMIKFGDV LTVNPLVVNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC	NiVG protein attachment glycoprotein Truncated Δ10
603	MGKGG IPSKVIKSY GTMDIKKINE GLLDSKILSA FNTVIALGGS IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKIHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLDRG DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCVAVST VGDPILNSTY WSGSLMMTRL AVKPKSNGGG YNQHQALALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCRLSM GIRPNSHYIL RSGLLKYNLS DGENPKVFI EISDQRLSIG SPSKIYDSLQ QPVFYQASFS WDTMIKFGDV LTVNPLVVNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC	NiVG protein attachment glycoprotein Truncated Δ15
604	MGSKVIKSY GTMDIKKINE GLLDSKILSA FNTVIALGGS IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKIHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLDRG DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCVAVST VGDPILNSTY WSGSLMMTRL AVKPKSNGGG YNQHQALALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCRLSM GIRPNSHYIL RSGLLKYNLS DGENPKVFI EISDQRLSIG SPSKIYDSLQ QPVFYQASFS WDTMIKFGDV LTVNPLVVNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV	NiVG protein attachment glycoprotein Truncated Δ20

	FKDNEILYRA QIASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC	
605	MGSYY GTMDIKKINE GLLDSKILSA FNTVIALLG IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLD RG DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCVST VGDPILNSTY WSGSLMMTRL AVKPKSNGGG YNQHLALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCRISM GIRPN SHYIL RSGLLKYNLS DGENPKVFI EISDQRLSIG SPSKIYDSLQ QPVFYQASFS WDTMIKFGDV LTVNPLVNVW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QIASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC	NiVG protein attachment glycoprotein Truncated Δ25
606	MGTMDIKKINE GLLDSKILSA FNTVIALLG IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLD RG DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCVST VGDPILNSTY WSGSLMMTRL AVKPKSNGGG YNQHLALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCRISM GIRPN SHYIL RSGLLKYNLS DGENPKVFI EISDQRLSIG SPSKIYDSLQ QPVFYQASFS WDTMIKFGDV LTVNPLVNVW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QIASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC	NiVG protein attachment glycoprotein Truncated Δ30
607	MKKINEGLDSKILSA FNTVIALLG IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLD RG DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCVST VGDPILNSTY WSGSLMMTRL AVKPKSNGGG YNQHLALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCRISM GIRPN SHYIL RSGLLKYNLS DGENPKVFI EISDQRLSIG SPSKIYDSLQ QPVFYQASFS WDTMIKFGDV LTVNPLVNVW RNNTVISRPG QSQCPRFNTC PAICAEVYN DAFLIDRINW ISAGVFLDSN ATAANPVFTV FKDNEILYRA QIASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QCT	NiVG protein attachment glycoprotein Truncated and mutated (E501A, W504A, Q530A, E533A) NiV G protein (Gc Δ 34)
608	MATQEVRLKC LLCGIIVLVL SLEGLGILHY EKLSKIGLVK GITRKYKIKS NPLTKDIVIK MIPNVSNVSK CTGTVMENYK SRLTGILSPI KGAIELYNNN THDLVGDVKL AGVVMAGIAI GIATAAQITA GVALYEAMKN ADNINKLKSS IESTNEAVVK LQETAETVY VLTALQDYIN TNLVPTIDQI	Hendra virus F protein Uniprot O89342 (with signal sequence)

	SCKQTEALD LALSKYLSDL LFVFGPNLQD PVSNSMTIQA ISQAFGGNYE TLLRTLGYAT EDFDDLLESD SIAGQIVYVD LSSYYIIVRV YFPILTEIQQ AYVQELLPVS FNNDNSEWIS IVPNFVLIRN TLISNIEVKY CLITKKSVC NQDYATPMTA SVRECLTGST DKCPRELVS SHVPRFALSG GVLFANCISV TCQCQTGRA ISQSGEQTLL MIDNTTCTTV VLGNIISLG KYLGSINYNES ESIAGVPPVY TDKVDISSQI SSMNQSLQQS KDYIKEAQKI LDTVNPSSLIS MLSMIILYVL SIAALCIGLI TFISFVIVEK KRGYSRLDD RQVRPVSNGD LYYIGT	
609	MMADSKLVSL NNNLSGKIKD QGKVIKYYG TMDIKKINDG LLDSKILGAF NTVIALLSGI IIIVMNIMII QNYTRTTDNQ ALIKESLQSV QQQIKALTDK IGTEIGPKVS LIDTSSTITI PANIGLLGSK ISQSTSSINE NVNDCKKFTL PPLKIHENI SCPNPLPFRE YRPISQGVSD LVGLPNQICL QKTTSTILKP RLISYTLFIN TREGVCITDP LLAVDNGFFA YSHLEKIGSC TRGIAKQRII GVGEVLDRGD KVPSMFMTNV WTPPNPSTIH HCSSTYHEDE YYTLCAVSHV GDPILNSTSW TESLSLIRLA VRPKSDSGDY NQKYIAITKV ERGKYDKVMP YGPGSIKQGD TLYFPAVGFL PRTEFQYNDS NCPIIHCKYS KAENCRLSMG VNSKSHYILR SGLLKYNLSL GGDIIILQFIE IADNRLTIGS PSKIYNLSLQ PVFYQASYSW DTMIKLGDVD TVDPLRVQWR NNSVISRPGQ SQCPRENVCP EVCWEGTYND AFLIDRLNWV SAGVYLSNQ TAENPVFAVF KDNEILYQVP LAEDDTNAQK TITDCFLEN VIWCISLVEI YDTGDSVIRP KLFVAKIPAQ CSES	Hendra virus G protein Uniprot O89343
610	MVVILDKRCY CNLLILILMI SECSVGILHY EKLSKIGLVK GVTRKYKIKS NPLTKDIVIK MIPNVSNMSQ CTGSVMENYK TRNLGILTPI KGALEIYKNN THDLVGDVRL AGVIMAGVAI GIATAAQITA GVALYEAMKN ADNINKLKSS IESTNEAVVK LQETAECTVY VLTALQDYIN TNLVPTIDKI SCKQTELSLD LALSKYLSDL LFVFGPNLQD PVSNSMTIQA ISQAFGGNYE TLLRTLGYAT EDFDDLLESD SITGQIIYVD LSSYYIIVRV YFPILTEIQQ AYIQELLPVS FNNDNSEWIS IVPNFILVRN TLISNIEIGF CLITKRSVIC NQDYATPMTN NMRECLTGST EKCPRELVS SHVPRFALSN GVLFANCISV TCQCQTGRA ISQSGEQTLL MIDNTTCPTA VLGNIISLG KYLGSVYNYS EGIAIGPPVF TDKVDISSQI SSMNQSLQQS KDYIKEAQL LDTVNPSSLIS MLSMIILYVL SIASLCIGLI TFISFIIIVEK KRNTGT	Nipah virus NiV-F F0 truncation (aa 525-544) (with signal sequence)
611	MVVILDKRCY CNLLILILMI SECSVGILHY EKLSKIGLVK GVTRKYKIKS NPLTKDIVIK MIPNVSNMSQ CTGSVMENYK TRNLGILTPI KGALEIYKNQ THDLVGDVRL AGVIMAGVAI GIATAAQITA GVALYEAMKN ADNINKLKSS IESTNEAVVK LQETAECTVY VLTALQDYIN TNLVPTIDKI SCKQTELSLD LALSKYLSDL LFVFGPNLQD PVSNSMTIQA ISQAFGGNYE TLLRTLGYAT EDFDDLLESD SITGQIIYVD LSSYYIIVRV YFPILTEIQQ AYIQELLPVS FNNDNSEWIS IVPNFILVRN	Nipah virus NiV-F F0 truncation (aa 525-544) AND mutation on N-linked glycosylation site (with signal sequence)

	TLISNIEIGF CLITKRSVIC NQDYATPMTN NMRECLTGST EKCPRELVS SHVPRFALSN GVLFANCISV TCQCQTTGRA ISQSGETLL MIDNTTCPTA VLG NVIISLG KYLGSVNYS EGIAIGPPVF TDKVDISSQI SSMNQSLQQS KDYIKEAQRL LDTVNPSLIS MLSMIILYVL SIASLCIGLI TFISFIIVEK KRNTGT	
599	MVVILDKRCY CNLLILILMI SECSVGILHY EKLSKIGLVK GVTRKYKIKS NPLTKDIVIK MIPNVSNMSQ CTGSVMENYK TRLNGILTPI KGALEIYKNN THDLVG DVRL AGVIMAGVAI GIATAAQITA GVALYEAMKN ADNINKLKSS IESTNEAVVK LQETAECTVY VLTALQDYIN TNLVPTIDKI SCKQTELSLD LALSKYLS DL LFVFGPNLQD PVSNSMTIQA ISQAFGGNYE TLLRTLGYAT EDFDDLLESD SITGQIIYVD LSSYYIIVRV YFPILTEIQQ AYIQELLPVS FNNDNSEWIS IVPNFILVRN TLISNIEIGF CLITKRSVIC NQDYATPMTN NMRECLTGST EKCPRELVS SHVPRFALSN GVLFANCISV TCQCQTTGRA ISQSGETLL MIDNTTCPTA VLG NVIISLG KYLGSVNYS EGIAIGPPVF TDKVDISSQI SSMNQSLQQS KDYIKEAQRL LDTVNPSLIS MLSMIILYVL SIASLCIGLI TFISFIIVEK KRNT	Truncated NiV fusion glycoprotein (FcDelta22) at cytoplasmic tail (with signal sequence)
612	MKKINEGLLDKILSA FNTVIALLGSI IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKIHCECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLDRG DEVPSLFMTN WSTPPNPNTV YHCSAVYNNE FYYVLCVAVST VGDPIILNSTY WSGSLMMTRL AVKPKSNGGG YNQHQALARS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCRSLM GIRPN SHYIL RSGLLKYNLS DGENPKVVFI EISDQRLSIG SPSKIYDSLQ QPVFYQASFS WDTMIKFGDV LTVNPLVNVN RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QCT	NiVG protein attachment glycoprotein Truncated (Gc Δ 34)
613	ILHY EKLSKIGLVK GVTRKYKIKS NPLTKDIVIK MIPNVSNMSQ CTGSVMENYK TRNLGILTPI KGALEIYKNN THDLVG DVRL AGVIMAGVAI GIATAAQITA GVALYEAMKN ADNINKLKSS IESTNEAVVK LQETAECTVY VLTALQDYIN TNLVPTIDKI SCKQTELSLD LALSKYLS DL LFVFGPNLQD PVSNSMTIQA ISQAFGGNYE TLLRTLGYAT EDFDDLLESD SITGQIIYVD LSSYYIIVRV YFPILTEIQQ AYIQELLPVS FNNDNSEWIS IVPNFILVRN TLISNIEIGF CLITKRSVIC NQDYATPMTN NMRECLTGST EKCPRELVS SHVPRFALSN GVLFANCISV TCQCQTTGRA ISQSGETLL MIDNTTCPTA VLGNVIISLG KYLGSVNYS EGIAGPPVF TDKVDISSQI SSMNQSLQQS KDYIKEAQRL LDTVNPSLIS MLSMIILYVL SIASLCIGLI TFISFIIVEK KRNT	Truncated mature NiV fusion glycoprotein (FcDelta22) at cytoplasmic tail
614	MSNKRRTVLIIISYTLFYLNNAAIVGFDFDKLNKIGVVQGRVLN YKIKGDPMTKDLVLKFIPIVNVNITECVREPLSRYNETVRLLLP IHNMLGLYLNNNTNAKMTGLMIAGVIMGGIAIGIATAAQITAGFA LYEAKNTENIQKLTD SIMKTQDSIDKLTDSVGTSLILNLKLT YINNQLVPNLELLSCRQNKIEFDLMLTKYLVDLMTVIGPNINNP VNKDMTIQSLSLFDGNYDIMMSELGYTPQDFLDLIESKSITGQ	gb:JQ001776:6129- 8166 Organism:Ceda r virus Strain Name:CG1a Protein Name:fusion glycoprotein Gene

	IIYVDMENLYVVIRTYLPTLIEVPDAQIYEFNKITMSSNGGEYL STIPNFILIRGNYSNIDVATCYMTKASVICNQDYSLPMSQNL SCYQGETEYCPVEAVIASHSPRFALTNGVIFANCINTICRCQDN GKTITQNNINQFVSMIDNSTCNDVMVDKFTIKVGKYMGRKDINN NIQIGPQIIIDKVDLSNEINKMNQSLKDSIFYLREAKRILDSVN ISLISPSVQLFLIIISVLSFIILLIIIVYLYCKSKHSYKYNKFI DDPDYYNDYKRERINGKASKSNNIYYVGD	Symbol:F (with signal sequence)
615	MALNKNMFSSFLGYLLVYATTVQSSIHYDSLKVGVKGLTYN YKIKGSPSTKLMVVKLIPNIDSVKNCTQKQYDEYKNLVRKALEP VKMAIDTMLNNVSGNNKYRFAGAIMAGVALGVATAATVTAGIA LHRSENENAQAIANMKSATQNTNEAVKQLQLANKQTLAVIDTIRG EINNNIIPVINQLSCDTIGLSVGIRLTQYYSEIITAFGPALQNP VNTRITIQAISSVFNGNFDELLKIMGYTSGDLYEILHSELIRGN IIDVDVDAGYIALEIEFPNLTLPVNAVQELMPISYNIDGDEWV TLVPRFVLTRTLLSNIDTSRCTITDSSVICDNDYALPMSHELI GCLQGDTSKAREKVVSSYVPKFALSDGLVYANCLNTICRCMDT DTPISQSLGATVSLLDNKRCVYQVGDVLSVGSYLGDEYNAD NVELGPPIVIDKIDIGNQLAGINQTLQEAEDYIEKSEEFKGVN PSIITLGSMLVLYIFMILIAIVSVIALVLSIKLTVKGNVVRQQF TYTQHVPSMENINYVSH	gb:NC_025352:5950-8712 Organism:Mojang virus Strain Name:Tongguan1 Protein Name:fusion protein Gene Symbol:F (with signal sequence)
616	MKKKTDNPTISKRGHNHSRGIKSRALLRETDNYSNGLIVENLVR NCHHPSKNNLNYTKTQKRDSTIPYRVEERKGYHYPKIKHLIDKSY KHIKRGKRRNGHNGNIITIIILLILILIKTQMSEGAIHYETLSKI GLIKGITREYKVGTPSSKDIVIKLIPNVTGLNKCTNISMENYK EQLDKILIPINNIIELYANSTKSAPGNARFAGV IAGVALGVAA AAQITAGIALHEARQNAERINLLKDSISATNNVAELQEATGGI VNVITGMQDYINTNLVPQIDKLQCSQIKTALDISLSQYYSEILT VFGPNLQNPVTTSMSTQAISSQSGFNNIDLLNLLGYTANDLLDL LESKSITGQITYINLEHYFMVIRVYYPIMTTISNAYVQELIKIS FNVGDSEWVSLVPSYILIRNSYLSNIDISECLITKNSVICRHDF AMPMSYTLKECLTGDTEKCPREAVVTSYVPRFAISGGVIYANCL STTCQCYQTGKVIADGSGQTLMMIDNQTCISIVRIEILISTGKY LGSQEYNTMHVSVGNPVFTDKLDITSQISNINQSIEQSKFYLDK SKAILDKINLNLIGSVPIISILFIIAILSLILSIITFVIVMIIIVR RYNKYTPLINSDPSSRRSTIQDVYIIPNPGEHSSIRSAARSIDRD RD	gb:NC_025256:6865-8853 Organism:Bat Paramyxovirus Eid_hel/GH-M74a/GHA/2009 Strain Name:BatPV/Eid_hel/GH-M74a/GHA/2009 Protein Name:fusion protein Gene Symbol:F (with signal sequence)
27	(GGGGGS) _n wherein n is 1 to 6	Peptide Linker
618	MPAENKKVRFENTTSKGGKIPSKVIKSYGGTMDIKKINEGLLDS KILSAFNTVIALLGSIIVIMNIMIIQNYTRSTDNQAVIKDALQ GIQQQIKGLADKIGTEIGPKVSLIDTSSTITIPANIGLLGSKIS QSTASINENNVNEKCKFTLPPLKIHECNISCPNPLPFREYRPQTE GVSNLVGLPNNICLQKTSNQILPKLISYTLPPVVGQSGTCITDP LLAMDEGYFAYSHLERIGSCSRGVSKQRIIGVGEVLDRGDEVPS LFMTNVWTPPNPNTVYHCSAVYNNEFYVLCVAVSTVGDPILNST YWSGSLMMTRLAVKPKSNGGGYNQHQLALRSIEKGRYDKVMPYG PSGIKQGD TLYFPAVGFLVRTEFKYNDSNCPITKCQYSKPENCR LSMGIRPNSHYILRSGLLKYNLSDGENPKVVFIEISDQRLSIGS PSKIYDSLQGPVFYQASFSDWTMIKFGDVLTVNPLVNVNRNNTV ISRPGQSQCPRFNTCPEICWEGVYNDAFLIDRINWISAGVFLDS NQTAENPVFTVFKDNEILYRAQLASEDTNAQKTITNCFLLKNKI WCISLVEIYDTGDNVIRPKLFAVKIPEQCT	gb:AF212302 Organism:Nipah virus Strain Name:UNKNOWN-AF212302 Protein Name:attachment glycoprotein Gene Symbol:G (Uniprot Q9IH62)
619	MLSQQLQKNYLDNSNQQGDKNMNPDKKLSVNFNPLELDKGQKDLN KSYVVKNNKYNVSNLLNESLHDIKFCIYCFSLIIITIIINIIT ISIVITRLKVHEENGMESPNLQSIQDSLSSLTNMINTETPRI	gb:JQ001776:8170-10275 Organism:Cedar virus Strain

	GILVTATSVTLSSSINYVGTKTNQLVNELKDYITKSCGFKVPEL KLHECNISCADPKISKAMYSTNAYAELAGPPKIFCKSVSKDPD FRLKQIDYVIPVQQDRSICMNNPLLDISDGFYIHYEGINSCK KSDSFKVLLSHGEIVDRGDYRPSLYLLSSHYHPYSMQVINCPV TCNQSSFVFCCHISNNTKTLDNSDYSSDEYYITYFNGIDRPKTK IPINNMTADNRYIHFTFSGGGGVCLGEEFIIPVTTVINTDVFT DYCESFNCSVQTGKSLKEICSESLRSPNTSSRYNLNGIMIISQN NMTDFKIQLNGITYNKLSFGSPGRLSKTLGQVLYYQSSMSWDY LKAGFVEKWKPFPTNWMNNTVISRPNQGNCPRYHKCPEICYGGT YNDIAPLDLGKDMYVSVILDSQLAENPEITVFNSTTILYKERV SKDELNTRSTTTSCFLFLDEPWCISVLETNRFNKGKSIRPEIYSY KIPKYC	Name:CG1a Protein Name:attachment glycoprotein Gene Symbol:G
620	MPQKTVEFINMNSPLERGVSTLSDKKTNLNQSKITKQGYFGLGSH SERNNKKQKNQNDHYMTVSTMILEILVVLGIMFNLIIVLTMVYYQ NDNINQRMAELTSNITVLNLNLNQLTNKIQREIIPRITLIDTAT TITIPSAITYILATLTTRISELLPSINQKCEFKTPTLVLNDCRI NCTPPLNPSDGVKMSSLATNLVAHGSPPCRNFSSVPTIYYRIP GLYNRTALDERCILNPRLTISSTKFAYVHSEYDKNCTRGFKYYE LMTFGEILEGPEKEPRMFSRSFYSPNAVNYHSCTPIVTVNEGY FLCLECTSSDPLYKANLSNSTFHLVILRHNKDEKIVSMPSFNLS TDQEYVQIIPAEGGGTAESGNLYFPCIGRLLHKRVTHTPLCKKSN CSRTDDESLKSYNQGSPQHQQVNVCLIRIRNAQRDNPTWDVIT VDLTNTYPGSRSRIFGFSFKPMLYQSSVSWHTLLQVAEITDLDK YQLDWLDTPIISRPGGSECPFGNYCPTVCWEGTYNDVYSLTPNN DLFVTVYLKSEQVAENPYFAIFSRDQILKEFPLDAWISSARTTT ISCFMFNNEIWCIAALEITRLNDDIIRPIYYSFWLPTDCRTPYP HTGKMTRVPLRSTYNY	gb:NC_025256:9117- 11015 Organism:Bat Paramyxovirus Eid_hel/GH- M74a/GHA/2009 Strain Name:BatPV/Eid_hel /GH- M74a/GHA/2009 Protein Name:glycoprotein Gene Symbol:G
621	MATNRDNTITSAEVSQEDKVKKYYGVETAEKVADSISGNKVFIL MNTLLILTGAIIITITLNTNLTAAKSQQNMLKIIQDDVNAKLEM FVNLDQLVKGEEKPKVSLINTAVSVSIPGQISNLQTKFLQKYVY LEESITKQCTCNPLSGIFPTSGPTYPTDKPDDDTTDDDQVDTT IKPIEYKPKDGCNRTGDHFTMEPGANFYTPVNLGPASSNSDECY TNPSFSIGSSIYMFSQEIRKTDCTAGEILSIQIVLGRIVDKGQQ GPQASPLLWAVPNPKIINSCAVAAGDEMGLVCSVTLTAAASGE PIPHMFDGFWLYKLEPDTEVVSYRITGYAYLLDKQYDSVFIGKG GGIQKGNLDLYFQMYGLSRNRQSFKALCEHGSCLGTGGGGYQVLC DRAVMSFGSEESLITNAYLKVNLDLASGKPVIIIGQTFPPSDSYKG SNGRMYTIGDKYGLYLAPSSWNRYLRFGITPDISVRSTTWLKSQ DPIMKILSTCTNTDRDMCPEICNTRGYQDIFPLSEDSEYYTYIG ITPNNGGTKNFVAVRDSGDHIASIDILQNYYSITSATISCFMYK DEIWCIAITEGKKQKDNQPRIYAHSYKIRQMCYNMKSATVTVGN AKNITIRRY	gb:NC_025352:8716- 11257 Organism:Mojang virus Strain Name:Tongguan1 Protein Name:attachment glycoprotein Gene Symbol:G
622	FNTVIALLGSIIVIMNIMI IQNYTRSTDN QAVIKDALQG IQQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKIHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLDRL DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCVAVST VGDPIINSTY WSGSLMMTRL AVKPKSNGGG YNQHQLALRS IEKGRYDKVM PYGPSGIKQG DTLYFPVAVGF LVRTEFKYND SNCPITKCQY SKPENCRLSM GIRPNSHYIL RSGLLKYNLS DGENPKVVFI EISDQRLSIG SPSKIYDSL QPVFYQASFS WDTMIKFGDV LTVNPLVVNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN	NivG protein attachment glycoprotein Without cytoplasmic tail Uniprot Q9IH62

	QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC	
623	FNTVIALLSI IIIVMNIMI QNYTRTTDNQ ALIKESLQSV QQQIKALTDK IGTEIGPKVS LIDTSSTITI PANIGLLGSK ISQSTSSINE NVNDKCKFTL PPLKIHECNI SCPNPLPFRE YRPISQGVSD LVGLPNQICL QKTTSTILKP RLISYTLPIN TREGVCITDP LLAVDNGFFA YSHLEKIGSC TRGIAKQRRI GVGEVLDRGD KVPSMFMNTV WTPPNPSTIH HCSSTYHEDF YYTLCAVSHV GDPILNSTSW TESLSLIRLA VRPKSDSGDY NQKYIAITKV ERGKYDKVMP YGPSGIKQGD TLYFPAVGFL PRTEFQYNDS NCPIIHCKYS KAENCRLSMG VNSKSHYILR SGLLKYNLSL GGDIIILQFIE IADNRLTIGS PSKIYNSLGQ PVFYQASYSW DTMIKLGDVD TVDPLRVQWR NNSVISRPGQ SQCPRFNVCP EVCWEGTYND AFLIDRLNWV SAGVYLSNQ TAENPVFAVF KDNEILYQVP LAEDDTNAQK TITDCFLEN VIWCISLVEI YDTGDSVIRP KLFPAVKIPAQ CSES	Hendra virus G protein Uniprot O89343 Without cytoplasmic tail
624	MVVILDKRCY CNLLILILMI SECSVG	Signal sequence
625	GGGGGS	Peptide linker
626	(GGGGS)n wherein n is 1 to 10	Peptide linker
627	GGGGGS	Peptide linker
628	PAENKKVR FENTTSKDGK IPSKVIKSY GTMDIKKINE GLLDSKILSA FNTVIALLSI IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKIHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLDRG DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCVST VGDPIINSTY WSGSLMTRL AVKPKSNGGG YNQHLALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGFL LVRTEFKYND SNCPITKCQY SKPENCRISM GIRPNSHYIL RSGLLKYNLS DGENPKVFEI EISDQRLSIG SPSKIYDSLQ QPVFYQASFS WDTMIKFGDV LTVNPLVNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRLNW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC	NiVG protein attachment glycoprotein (602 aa) Without N-terminal methionine
629	KVR FENTTSKDGK IPSKVIKSY GTMDIKKINE GLLDSKILSA FNTVIALLSI IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKIHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLDRG DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCVST VGDPIINSTY WSGSLMTRL AVKPKSNGGG YNQHLALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGFL LVRTEFKYND SNCPITKCQY SKPENCRISM GIRPNSHYIL RSGLLKYNLS DGENPKVFEI EISDQRLSIG SPSKIYDSLQ QPVFYQASFS WDTMIKFGDV LTVNPLVNW	NiVG protein attachment glycoprotein Truncated Δ5 Without N-terminal methionine

	RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC	
630	NTTSDKGK IPSKVIKSYY GTMDIKKINE GLLDSKILSA FNTVIALGGS IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKIHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLD RG DEVPSLFMTN VWTPPNPNTV YHCSAVYNNE FYYVLCVAVST VGDPIILNSTY WSGSLMMTRL AVKPKSNGGG YNQHQLALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCRLSM GIRPNSHYIL RSGLLKYNLS DGENPKVVFI EISDQRLSIG SPSKIYDSL QPVFYQASFS WDTMIKFGDV LTVNPLVVNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC	NiVG protein attachment glycoprotein Truncated Δ10 Without N-terminal methionine
631	KGK IPSKVIKSYY GTMDIKKINE GLLDSKILSA FNTVIALGGS IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKIHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLD RG DEVPSLFMTN VWTPPNPNTV YHCSAVYNNE FYYVLCVAVST VGDPIILNSTY WSGSLMMTRL AVKPKSNGGG YNQHQLALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCRLSM GIRPNSHYIL RSGLLKYNLS DGENPKVVFI EISDQRLSIG SPSKIYDSL QPVFYQASFS WDTMIKFGDV LTVNPLVVNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC	NiVG protein attachment glycoprotein Truncated Δ15 Without N-terminal methionine
632	SKVIKSYY GTMDIKKINE GLLDSKILSA FNTVIALGGS IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKIHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLD RG DEVPSLFMTN VWTPPNPNTV YHCSAVYNNE FYYVLCVAVST VGDPIILNSTY WSGSLMMTRL AVKPKSNGGG YNQHQLALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCRLSM GIRPNSHYIL RSGLLKYNLS DGENPKVVFI EISDQRLSIG SPSKIYDSL QPVFYQASFS WDTMIKFGDV LTVNPLVVNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC	NiVG protein attachment glycoprotein Truncated Δ20 Without N-terminal methionine
633	SYT GTMDIKKINE GLLDSKILSA FNTVIALGGS IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKIHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD	NiVG protein attachment glycoprotein

	PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLDRG DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCVAVST VGDPILNSTY WSGSLMMTRL AVKPKSNGGG YNQHLALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCRLSM GIRPNSHYIL RSGLLKYNLS DGENPKVVFI EISDQRLSIG SPSKIYDSLQ QPVFYQASFS WDTMIKFGDV LTVNPLVNVW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC	Truncated Δ25 Without N-terminal methionine
634	TMDIKKINE GLLDISKILSA FNTVIALLGSI IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKIHENI ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLDRG DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCVAVST VGDPILNSTY WSGSLMMTRL AVKPKSNGGG YNQHLALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCRLSM GIRPNSHYIL RSGLLKYNLS DGENPKVVFI EISDQRLSIG SPSKIYDSLQ QPVFYQASFS WDTMIKFGDV LTVNPLVNVW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC	NiV protein attachment glycoprotein Truncated Δ30 Without N-terminal methionine
635	KKINEGLLDISKILSA FNTVIALLGSI IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKIHENI ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLDRG DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCVAVST VGDPILNSTY WSGSLMMTRL AVKPKSNGGG YNQHLALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCRLSM GIRPNSHYIL RSGLLKYNLS DGENPKVVFI EISDQRLSIG SPSKIYDSLQ QPVFYQASFS WDTMIKFGDV LTVNPLVNVW RNNTVISRPG QSQCPRFNTC PAICAEGVYN DAFLIDRINW ISAGVFLDSN ATAANPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QCT	NiV protein attachment glycoprotein Truncated and mutated (E501 A, W504A, Q530A, E533A) NiV G protein (Gc Δ 34) Without N-terminal methionine
636	MADSKLVSL NNNLSGKIKD QGKVIKYYG TMDIKKINDG LLDISKILGAF NTVIALLGSI IIIVMNIMII QNYTRTTDNQ ALIKESLQSV QQQIKALTDK IGTEIGPKVS LIDTSSTITI PANIGLLGSK ISQSTSSINE NVNDKCKFTL PPLKIHENI SCPNPLPFRE YRPISQGVSD LVGLPNQICL QKTTSTILKP RLISYTLPIN TREGVCITDP LLAVDNGFFA YSHLEKIGSC TRGIAKQRII GVGEVLDRGD KVPSPMFTNV WTPPNPSTIH HCSSTYHEDF YYTLCVAVSHV GDPILNSTSW TESLSLIRLA VRPKSDSGDY NQKYIAITKV ERGKYDKVMP YGPSGIKQGD TLYFPAVGFL PRTEFQYND NCPIIHCKYS KAENCRLSMG VNSKSHYILR SGLLKYNLSL GGDIILOFIE IADNRLTIGS	Hendra virus G protein Uniprot O89343 Without N- terminal methionine

	PSKIYNSLGQ PVFYQASYSW DTMIKLGDVD TVDPLRVQWR NNSVISRPGQ SQCPRENVCP EVCWEGTYND AFLIDRLNWV SAGVYLNSNQ TAENPVFAVF KDNEILYQVP LAEDDTNAQK TITDCFLLEN VIWCISLVEI YDTGDSVIRP KLFVAVKIPAQ CSES	
637	KKINEGLLDKILSA FNTVIALLS IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKIHEN ICPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLDRG DEVPSLFMTN VWTPPNPNTV YHCSAVYNNE FYYVLCVAVST VGDPILNSTY WSGSLMMTRL AVKPKSNGGG YNQHLALRS IEKGRYDKVM PYGPSGIKQG DTLYFFAVGF LVRTEFKYND SNCPITKCQY SKPENCRLSM GIRPNSHYIL RSGLLKYNLS DGENPKVFFI EISDQRLSIG SPSKIYDSLQ QPVFYQASFS WDTMIKFGDV LTVNPLVNVW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QCT	NiVG protein attachment glycoprotein Truncated (Gc Δ 34) Without N-terminal methionine
638	LSQLQKNYLDNSNQGDKNMNPDKLSVNFNPLELDKGQKDLNK SYYVKNKNYNVSNLLNESLHDIKFCIYCFISLLIIITIINIITI SIVITRLKVHEENNGMESPNLQSIQDSLSSLTNMINTEITPRIG ILVTATSVTLSSSINYVGTKTNQLVNELKDYITKSCGFKVPELK LHECNISCADPKISKSAMYSTNAYAEAGPPKIFCKSVSKDPDF RLKQIDYVIPVQQDRSICMNNPLLDISDGFYIHYEGINSCKK SDSFVKVLSHGEIVDRGDYRPSLYLLSSHYHPYSMQVINCPVT CNQSSVFVCHISNNTKTLDNSDYSSDEYYITYFNGIDRPKTKKI PINNMTADNRYIHFTFSGGGGVCLGEEFIIPVTTVINTDVFTHD YCESFNCSVQTGKSLKEICSESLRSPTNSSRYNLNGIMIISQNN MTDFKIQNLNGITYNKLSFGSPGRLSKTLGQVLYYQSSMSWDTYL KAGFVEKWKPFPTPNWMNNTVISRPNQGNCPRYHKCPEICYGGTY NDIAPLDLGKDMYVSVILDSQLAENPEITVFNSTTILYKERS KDELNTRSTTTSCFLFLDEPWCISVLETNRFNGKSIRPEIYSYK IPKYC	gb:JQ001776:8170- 10275 Organism:Ced ar virus Strain Name:CG1a Protein Name:attachment glycoprotein Gene Symbol:G Without N- terminal methionine
639	PQKTVEFINMNSPLERGVSTLSDDKTLNQSKITKQGYFGLGSHS ERNWKKQKNQNDHYMTVSTMLEILVVLGIMFNLIIVLTMVYYQN DNINQRMALTSNITVLNLLNLQLTNKIQREIIPRITLIDTATT ITIPSAITYILATLTTRISELLPSINQKCEFPTPLVLNDCRIN CTPPLNPSDGVKMSSLATNLVAHGSPPCRNFSVPTIYYYRIPG LYNRTALDERCILNPRLTISSTKFAYVHSEYDKNCTRGFKYYEL MTFGEILEGPEKEPRMFSRSFYSPNAVNYHSCPTIVTVNEGYF LLECTSSDPLYKANLSNSTFHLVILRHNKDEKIVSMPSFNLST DQEYVQII PAEGGGTAESGNLYFPCIGRLLHKRVTPLCKKSNC SRTDDESLKSYNQGSPPQHVVNCLIRIRNAQRDNPTWDVITV DLTNTYPGSRSRIFGSFSKPMYQSSVSWHTLLQVAEITDLKY QLDWLDTPIYISRPGGSECPFGNYCPTVCWEGTYNDVYSLTPNND LFVTVYLKSEQVAENPYFAIFSRDQILKEFPLDAWISSARTTTI SCFMFNNEIWCIAALEITRLNDDIIRPIIYYSFWLPTDCRTPYPH TGKMTRVPLRSTYNY	gb:NC_025256:9117- 11015 Organism:Bat Paramyxovirus Eid_hel/GH- M74a/GHA/2009 Str ain Name:BatPV/Eid_hel /GH- M74a/GHA/2009 Pr otein Name:glycoprotein Gene Symbol:G Without N-terminal methionine
640	ATNRDNTITSAEVSQEDKVKYGYGVETAEKVADSI SGNKVFILM NTLLILTGAIIITITLNTNLTAAKSQQNMLKI IQDDVNAKLEMF VNLDQLVKGEEKPKVSLINTAVSVSIPGQISNLQTKFLQKYVYL EESITKQCTCNPLSGIFPTSGPTYPPTDKPDDDTTDDDKVDTTI	gb:NC_025352:8716- 11257 Organism:Moj iang virus Strain Name:Tongguan1 Pr

	KPIEYKPKPDGCNRTGDHFTMEFGANFYTVPNLGPASSNSDECYT NPSFSIGSSIYMFSEIRKTDCTAGEILSIQIVLGRIVDKGQQG PQASPLLVWAVPNPKI INSCAVAAGDEMGWVLC SVTLTAASGEP IPHMFDGFWLYKLEPDTEVVSYRITGYAYLLDKQYDSVFIGKGG GIQKGNLDYFQMYGLSRNRQSF KALCEHGSCLGTGGGGYQVLCD RAVMSFGSEESLITNAYLKVN DLASGKPVII GQTFPPSDSYKGS NGRMYTIGDKYGLYLAPSSWNRYLRFGITPDISVRSTTWLKSQD PIMKILSTCTNTDRDMCPEICNTRGYQDIFPLSEDSEYYTYIGI TPNNGGKTKNFVAVRDS DGHASIDILQNYYSITSATISCFMYKD EIWCIAITEGKKQKDN PQRIYAHSYKIRQMCYNMKSATVTVGNA KNITIRRY	otein Name:attachment glycoprotein Gene Symbol:G Without N- terminal methionine
641	DFDKLNKIGVVQGRVLNYKIKGDPMTKDLVLKFIPNIVNITECV REPLSRYNETVRRLLLPIHNMLGLYLNNTNAKMTGLMIAGVIMG GIAIGIATAAQITAGFALYEAKNTENIQKLTDSIMKTQDSIDK LTDSVGTSLILNLKLTQYINNQLVPLELLSCRQNKIEFDLMLT KYLVDLMTVIGPNINNPVNKDMTIQSLSLFDGNYDIMMSELGY TPQDFLDLIESKSITGQIIYVDMENLYVVIRTYLPTLIEVPDAQ IYEFNKITMSSNGGEYLSTIPNFILIRGNYSNIDVATCYMTKA SVICNQDYSLPMSQNLRSYQGETEYCPVEAVIASHSPRFALTN GVIFANCINTICRCQDNGKITQINQFVSMIDNSTCNDVMVDK FTIKVGKYMGRKDINNINIQIGPQIIIDKVDLSNEINKMNQSLK DSIFYLREAKRILDSVNISLISPSVQLFLIIISVLSFIIILLIII VYLYCKSKHSYKYNKFIDDPDYNDYKRERINGKASKSNIIYYV GD	gb:JQ001776:6129- 8166 Organism:Ceda r virus Strain Name:CG1a Protein Name:fusion glycoprotein Gene Symbol:F (without signal sequence)
642	SRALLRETDNYSNGLIVENLVRNCHHPSKNNLN YTKTQKRDSTI PYRVEERKGHPKIKHLIDKSYKH IKRGKRRNGHNGNIITII LL LILILKTQMSEGAIHYETLSKIGLIKGITREYKVKGT PSSKDIV IKLIIPNVTGLNKCTNISMENYKEQLDKILIPINNII ELYANSTK SAPGNARFAGVIIAGVALGVAAA AQITAGIALHEARQNAERINL LKDSISATNNAVAELQEATGGIVNVITGMQDYINTNLVPQIDKL QCSQIKTALDISLSQYYSEILTVFGPNLQNPVTTSMISQAISQS FGGNIDLLLNLGYTANDLLDLLESKSITGQITYINLEHYFMVI RVYYPIMTTISNAYVQELIKISFNVDGSEWVSLVPSYILIRNSY LSNIDI SECLITKNSVICRHFAMPMSYTLKECLTGDTEKCPRE AVVTSYVPRFAISGGVIYANCLSTTCQCYQTGKVIAQDGSQTL M MIDNQTC SIVRIEEILISTGKYLGSQEYNTMHVSVGNPVFTDKL DITSQISNINQSIEQSKFYLDKSKA ILDKINLNLIGSVPI SILF IIAILSLILSIITFVIVMIIVRRYNYKTPLINSDPSSRRSTIQD VYIIPNPGEHSIRSARSIDRDRD	gb:NC_025256:6865- 8853 Organism:Bat Paramyxovirus Eid_hel/GH- M74a/GHA/2009 Str ain Name:BatPV/Eid_hel /GH- M74a/GHA/2009 Pr otein Name:fusion protein Gene Symbol:F (without signal sequence)
643	ILHY EKLSKIGLVK GITRKYKIKS NPLTKDIVIK MIPNVSNVSK CTGTVMENYK SRLTGILSPI KGAIELYNNN THDLVGDKVL AGVVMAGIAI GIATAAQITA GVALYEAMKN ADNINKLKSS IESTNEAVVK LQETAECTVY VLTALQDYIN TNLVPTIDQI SCKQTEALD LALSKYLSDL LFVFGPNLQD PVSNSMTIQA ISQAFGGNYE TLLRTLGYAT EDFDDLLESD SIAGQIVYVD LSSYYIIIVRV YFPILTEIQQ AYVQELLPVS FNNDNSEWIS IVPNFVLIRN TLISNIEVKY CLITKKS VIC NQDYATPMTA SVRECLTGST DKCPREL VVS SHVPRFALSG GVLFANCISV TCQCQT TGRA ISQSGEQTLL MIDNTTCTTV VLGNIISLG KYLGSINYN S ESIAVGPPVY	Hendra virus F protein Uniprot O89342 (without signal sequence)

	TDKVDISSQI SSMNQSLQQS KDYIKEAQKI LDTVNPSLIS MLSMIILYVL SIAALCIGLI TFISFVIVEK KRGYSRLDD RQVRPVSNGD LYYIGT	
644	IHYDSLKVGVIKGLTYNYKIKGSPSTKLMVVKLIPNIDSVKNC TQKQYDEYKNLVRKALEPVKMAIDTMLNNVKSNNKYRFAGAIM AGVALGVATAATVTAGIALHRSNENAQAIAANMKSIAIQNTNEAVK QLQLANKQTLAVIDTIRGEINNNIIPVINQLSCDTIGLSVGIRL TQYYSEIITAFGPALQNPVNTRITIQAISSVFNGNFDELLKIMG YTSGLDLYEILHSELIRGNIIDVDVDAGYIALEIEFPNLTLVPNA VVQELMPISYNIDGDEWVTLVPRFVLTRTTLLSNIDTSRCTITD SSVICDNDYALPMSHELIGCLQGDTSKCAREKVVSYPKFAIS DGLVYANCLNTICRCMDTDTPISSQLGATVSLLDNKRCSVYQVG DVLISVGSYLGDGEYNADNVELGPPIVIDKIDIGNQLAGINQTL QEADYIEKSEEFKGVNPSIITLGSMMVLYIFMILIAIVSVIA LVLSIKLTVKGNVVRQQFTYTQHVPSMENINYVSH	gb:NC_025352:5950- 8712 Organism:Moji ang virus Strain Name:Tongguan1 Pr otein Name:fusion protein Gene Symbol:F (without signal sequence)
645	(GGGS)n wherein n is 1 to 10	Peptide linker
646	GGGSGGGSGGGGS	Peptide linker
647	TTAASGSSGGSSSGA	Peptide linker
648	GSTSGSGKPGSGEGSTKG	Peptide linker

Table 8. HCDRs using the Chothia Numbering Scheme

CD8 Binder	H-CDR1		H-CDR2		H-CDR3	
	Sequence	SEQ ID NO:	Sequence	SEQ ID NO:	Sequence	SEQ ID NO:
1	GGTFSSY	653	DPSDGN	701	ERAAAGYYYYMDV	148
2	GGTFNTY	654	DPSSGG	702	EHAAGTYYYYYMDV	149
3	GGTFSSY	653	DPSGGN	703	ERAAAGYYYYMDV	148
4	GYFTGY	655	NPNGGD	704	EGYYYYGMDV	649
5	GYFTDY	656	NPNSGG	705	EGDYGGMDA	150
6	GYTFTRY	657	NPNDGS	706	ERGGMPDY	151
7	GYFTSY	658	NPNSGG	705	GHGIPKY	152
8	GYFTSY	658	NPNSGN	707	VRSGSPQH	153
9	GHTFSRH	659	NPNSGN	707	GGPWIVDAFDI	154
10	GYFTSY	658	SAHNGV	708	GIAVAGTDY	650
11	GGTFSNT	660	NPSGGS	709	EATWGPYYYYMDV	155
12	GYTFTRS	661	SPYNGN	710	NKDGLQH	156
13	GDTFTGY	662	NPNSGD	711	DAKRVGYYYYMDV	157
14	GYTFTRY	657	NPNSGG	705	LVGGSPDY	158
15	GYFTNY	663	NPNSGG	705	GAMVDY	159
16	GGTFSNT	660	NPSDGD	712	GNVVGSYGGMDV	160
17	GYFTNY	663	NPNSGD	711	DSRGDWYFDL	161
18	GYGFTRY	664	DPSGGS	713	HGGRGLADY	162
19	GYFTSR	665	DPKSGD	714	LKELSSILDAFDI	163
20	GYFTSY	658	NPGAGS	715	ERFGTYGGYYMDV	164
21	GFTFSNS	666	SGDGGT	716	VIGEMVDDAFDL	165
22	GYFTGY	655	NPNSGD	711	ERLFGTYGGYYMDV	166
23	GYFTTY	667	IPIFGT	717	ADGELTDY	167
24	GFTFSSY	668	-GTGGG	718	HHLPAHYGGYYMDV	168
25	GGTFSRY	669	NPNSGD	711	DVPAGRYGGYYMDV	169
26	GNTFTSY	670	NPSDGS	719	DRGVGRYYGGYYMDV	170

27	GGTFSRY	669	NPSDGS	719	DSRYGRYYYYMDV	171
28	GGTFSNY	671	NPNGGS	720	EIVVGPYYYYMDV	172
29	GGTFTRY	672	NPNSGD	711	GMVRGPYYYYMDV	173
30	GGTFSSY	653	NPSGGS	709	EGVTGPYYYYMDV	174
31	GGTFSRF	673	NPSDGS	719	DAAAGTRYYYYGMDV	175
32	GGTFSSH	674	NPSGGS	709	ELYSSTYYYYMDV	176
33	GGTFSSY	653	NPNTGG	721	ALYSGPYYYYMDV	177
34	GFTFSNS	666	SGSGGS	722	EHAAGTYYYYMDV	149
35	GGTFGSY	675	SGYNGD	723	DSLVRGYYYYMDV	178
36	GYIFTDY	676	SADNGN	724	RSELDY	179
37	GYTFTSY	658	SPNSGA	725	GDDNDY	180
38	GYTFTSY	658	NPNSGN	707	GEEVDY	181
39	GYTFTSY	658	NPSGGS	709	GRRVPDY	182
40	GYTFTDY	656	NPKSGA	726	GKVTTDY	183
41	GFTFSSF	677	SESGDS	727	GRELIEY	184
42	GFTFDDY	678	-GTGGG	718	VYDFPDV	185
43	GYTFTDS	679	NPSNGD	728	STYSHIDY	186
44	GYTFTNY	663	SPSDGS	729	EDSSGFDY	187
45	GYTFMNY	680	NPSGGS	709	DQGGGFDY	188
46	GYTFTSY	658	DPEDGE	730	DQGWGMDV	189
47	GYTFTSY	658	NPKSGR	731	LTEGIPDY	190
48	GYTLNDY	681	NPGGGS	732	DRYGPFDY	191
49	GYTFTSY	658	NPKTGD	733	LVAGGAPDY	192
50	GYTFTGY	655	DPSDGY	734	DGFTGDIAY	193
51	GYTFTGY	655	NPNSGG	705	VDDSSSPDY	194
52	GYTFTGY	655	MPISGT	735	GPDGTEVDY	195
52	GYTFTNH	682	NPNSGN	707	SESGSDLDY	196
54	GYTFTNY	663	SPTSGD	736	EVEIEGYMDV	197
55	GYTFTSY	658	NPNSGD	711	DLDDDWYMDV	198
56	GYTFTSY	658	DPSGDI	737	DSTTWDAFDI	199
57	GYTFTDY	656	NPNSGG	705	VLVGSGSPDY	200
58	GYTFTEN	683	ETSGGS	738	AAAAGLDFQH	201
59	GYTFASY	684	NPNSGG	705	ANSWDADY	202
60	GFTFSNS	666	SGSGVT	739	EHSSSWYTFDY	203
61	GYTFTAY	685	NPNSGG	705	DDDSSGYLDY	204
62	GYTFTNY	663	NPSGGS	709	ASGDYMDLIDYMDY	205
63	GYTFTDY	656	NPDSGG	740	VGSSGYLAPTH	206
64	GYPFTDY	686	NPNSGN	707	VRGDGYNLGDY	207
65	GYTFSDY	687	NPNSGG	705	DVDTAMGAGDY	208
66	GYTFTDY	656	NPSGGS	709	VARWGYGDYPDY	209
67	GDTFTTH	688	SPSDGS	729	DRNGDYYYGMDV	210
68	GDTFTNY	689	NPISGG	741	EGLGSSWYVLDY	211
69	GYTFTSY	658	SADNGD	742	DGSHYGYGMDV	212
70	GYTFTSY	658	SPIYGT	743	PGPEGYYYGMDV	213
71	GYTFTDN	690	NPNSGN	707	YHWDYGDYRFDY	214
72	GYTFTSY	658	NPNSGN	707	VEIDYGDSPPDY	215
73	GGTSSSY	691	NPSDGD	712	GAEWELRYAFDI	216
74	GYTFTTY	667	NPSGGT	744	ETYYGLYYYGMDV	217
75	GYTFTSY	658	NPKSGN	745	APSLRGYSYGPDY	218
76	GGTFTSY	692	NPSGGS	709	DRQERYYYYYMDV	219
77	GYTFTSY	658	NPSDGS	719	DRSYGDYYYGMDV	220
78	GGTFTSY	692	NPGGGN	746	EVFSENNYYYMDV	221
79	GYTFTSY	658	NPSDGS	719	EWDYTHYYYGMDV	222

80	GNTFTSH	693	DPEDGE	730	GDSSGYQQYYFDY	223
81	GYTFTSY	658	TPVFGI	747	GSWDSSSWYIPEY	224
82	GFTFSDY	694	NPRGGS	748	LVWGGAYYYYMDV	225
83	GYTFTSY	658	NPNNGD	704	PVFSGSYYWYFDP	226
84	GYTFTSY	658	NPSGGG	749	DQAVAGPYYYGMDV	227
85	GGTFSSY	653	NPSGSN	750	DRWLAGPYYYGMDV	228
86	GYMFTGH	695	IPIFGT	717	VMGPVDYYYGMDV	229
87	GYIFSNY	696	NPSDGS	719	DLGPFGSYYYMDV	230
88	GFTFSSY	668	NGDGDD	751	EGVVVPPYYYMDV	231
89	GYTFTTY	667	DPNSGD	752	SSGWSRYYYMDV	232
90	GSTFTNY	697	NPSGGS	709	DNGMTTGYYYMDV	233
91	GYTFTSY	658	NPSGGS	709	DRAMVTGYYYGMDV	234
92	GYTFTSY	658	NPSDGN	753	DRGYGDRGYYYGMDV	235
93	GGTLSSY	698	NTYNGN	754	SPKATADYYYMDV	236
94	GYTFTSY	658	NPSDGI	755	STVTPSYYYGMDV	237
95	GYTFTSH	699	NPRDGD	756	EPVAGTGYYYGMDV	238
96	GGTFNSY	700	NPNSGN	707	DNLAGFWSYYYGMDV	239
97	GRTFSGY	1064	SRGGLS	1065	DRSDLYEITAASNIDS	1063

Table 9. LCDRs using the Chothia Numbering Scheme

CD8 Binder	L-CDR1		L-CDR2		L-CDR3	
	Sequence	SEQ ID NO:	Sequence	SEQ ID NO:	Sequence	SEQ ID NO:
1	RASQSISSYLN	240	AASSLQS	294	QQSYSTPLT	333
2	QASQDISNYLN	241	AASSLQS	294	QQSYSNLVS	334
3	RASQSISSYLN	240	AASSLQS	294	QQSYSTPLT	333
4	RSSQSLLSHNGYNYLD	242	LGSNRAS	295	MQALQTPFT	335
5	RSSQSLLSHNGYNYLD	242	LGSNRAS	295	MQGLQTPHT	336
6	RASQISRNLN	243	KASNLKG	296	QQTYSAPL	337
7	RSSQSLLSHNGYNYLD	242	LGSNRAS	295	MQLQTPLT	338
8	RASQSVASDLA	244	GASTRAT	297	QQYGDSPGS	339
9	QASQDIGNYLN	245	AASTLQR	298	QQANSFPPT	340
10	RASQSISTHLA	246	GASTRAT	297	QQYGNSRT	341
11	RASQTISNYLN	247	AASTLQS	299	QQSYSTPPT	342
12	RASQGIRNDLG	248	DASTLQS	300	QQSYSSPYT	343
13	RASQISNYLN	249	AASSLQS	294	QQSYSTPYT	344
14	RSSQSLLSHNGYNYLD	242	LGSNRAS	295	MQGAHWPPPT	345
15	RASQGISDSLA	250	GASSLRS	301	QQSYRTPYT	346
16	RASQISNYLN	249	AASSLQS	294	QESFTTQWT	347
17	QASQDIHNYLN	251	DASNLET	302	QQANSFPPT	340
18	QASQDISNYLN	241	SASSLQS	303	QQRSNWPLYT	348
19	RASQSIDWLA	252	AASSLQT	304	QQAISFPIT	349
20	QASQDISNYLN	241	SASTLQS	305	QQSYSPFT	350
21	RASQSISTWLA	253	AASTLQS	299	QQAISFPLT	351
22	RASQISNYLN	249	AASTLQS	299	QQSYTFPIT	352
23	RSSQSLLSHNGYNYLD	242	DASHLET	306	QQYYSYPPT	353
24	QASQDISNYLN	241	AASTLHS	307	QQSYSAPLT	354
25	QASQDISNYLN	241	AASTLQS	299	QQSFSTFYT	355
26	QASQDISNYLN	241	AASTLQS	299	QQSYSIPFT	356
27	RASQSINRFLN	254	AASSLQN	308	QQSYSTPYT	344
28	RASQSISSYLN	240	AASSLQS	294	QQSYSTPLT	333

29	QASQDISNYLN	241	AASTLQS	299	QQSYSTPIT	357
30	RASQSVSTYLN	255	AASSLQS	294	QQSYTIPST	358
31	QASQDIACYLN	256	AASSLQS	294	QQSYSAPPT	359
32	QASQGITNYLN	257	GASSLQS	309	QQSYSTPWT	360
33	RASQSISSYLN	240	AASSLQS	294	QQSYSTPLT	333
34	QASQDIHNYLN	251	AASTLQS	299	QQSYTTPLT	361
35	QASQDISNYLN	241	SAFSLQS	310	QQSYSAPIT	362
36	RASQSISSYLN	240	SASNLQS	311	QQRSNWPPVT	363
37	QANQDISNFLE	258	DASSLES	312	QQSYSIPIT	364
38	RASQGISNNLN	259	EASTLES	313	QQSYSTPLT	333
39	RSSQSLLSHNGYNYLD	242	GASTLET	314	MQGLQPPGT	365
40	RASQSISRSLV	260	AASTLQT	315	QQSYNHFRIT	366
41	QASQDISNYLN	241	DASNLET	302	QRSDSTPLT	367
42	QASHDISKSLN	261	GASTLQS	316	QQLNSYPRT	368
43	RASQDIGAYLA	262	AASSLQS	294	QQSYSIPYT	369
44	RASQSISSYLA	263	AASSLQS	294	QQSYSTPYT	344
45	RASQGIRSYLA	264	GASNLET	317	QQSYSTPYT	344
46	RASQSISSYLN	240	AASSLQS	294	QQTYSTPYT	370
47	RASQNIGTWLA	265	AASTLQS	299	QQSYSTPQT	371
48	RASQTISYYLN	266	AASTLQS	299	QQSYRTPYT	346
49	RSSQSLLSHNGYNYLD	242	MGSNRAS	318	MQGTHWPT	372
50	RASQNINNYLN	267	GASSLQS	309	QQTFSLPYT	373
51	RASQTISTYLN	268	DASNLET	302	QQSYSTPYT	344
52	RASRGIGNDLA	269	DASTLET	319	QQGYNMPLT	374
52	RASQTIGNYVN	270	GASNLHT	320	QQTYSAPLT	375
54	RASQFIGSWLA	271	AASTLQS	299	QQSYSFPWT	376
55	RASQSISSWMA	272	DASNLET	302	QQTYSTPYI	377
56	RASQGISNNLN	259	DASNLET	302	QQSYSSPWT	378
57	KSSQSVLYSSNNKNYLA	273	WASTRES	321	QQYASAPRT	379
58	RASQSISSYLN	240	KTSSLES	322	QQSFTIPYT	380
59	RVSQGISSYLN	274	GASSLQS	309	QQSYSTPLT	333
60	RASQSIDWLA	252	DASNLET	302	QQSYSTPLT	333
61	RASQGISNYLA	275	SASNLQS	311	QQTyrTPPT	381
62	RASQSIRNYLT	276	SASNLQS	311	QQSYSTPLT	333
63	RASQNIRLYLN	277	AASTLQS	299	QQLTTPFT	382
64	QASQDIRKFLN	278	AASSLQS	294	QQLNGYPGT	383
65	RASQSISSYLN	240	TASNLQS	323	QQSYSLPLT	384
66	QASQDISNYLS	279	DASNLQS	324	QQTYYTPRT	385
67	RASQNVRSWLA	280	AASSLQS	294	QQSYNTPYT	386
68	RASQGIGNDLG	281	AASSLQS	294	QQSYAPPPT	387
69	RASQISNWLA	282	GASNLET	317	QQSYSTPPT	342
70	RSSQSLLSHNGYNYLD	242	LGSNRAS	295	MQGLQTPLT	388
71	RASQSISSYLN	240	LASSLQS	325	QQSDSIPVT	389
72	QASQDISNYLN	241	STSSLQS	326	QQSYSTPYN	390
73	RASESIGSWLA	283	AASSLQS	294	QQSYSTPYT	344
74	RASQISNYLN	249	AASSLQR	327	QQSYSTPLT	333
75	RASQSVTSNYLA	284	GASTRAT	297	QHYGSSPA	391
76	RASQSISSYLN	240	AASSLQS	294	QQSYSTPLT	333
77	RASQGISSYLA	285	AASTLQS	299	QQSYSTPPT	342
78	RASQDIGNYLN	286	AASSLQS	294	QQAYTYPYT	392
79	QASQDISNYLN	241	GASSLQS	309	QQSYTTPNT	393
80	RASQGISNYLA	275	AASTLQS	299	QQSYSTPYT	344
81	RASQGISNGLS	287	DASNLET	302	QQSYSTPFT	394

82	RASQNIRNYLN	288	GASSLQS	309	QQSYSTPLT	333
83	QASLDINNYLN	289	KASSLES	328	QQSYSMPLT	395
84	QASQDISNYLN	241	AASSLQG	329	QQSYTTPWT	396
85	QASQDISNYLN	241	AASSLQS	294	QQSYSSPLT	397
86	QASQDISNYLN	241	KASSLES	328	QQSYSDPLT	398
87	QASQDISNYLN	241	GASTLQS	316	QQSYSAPIT	362
88	RASQSISNYLN	249	AASNLQS	330	QQSYTTPLT	361
89	RASQNIGNYLN	290	AASTLQS	299	QQSYSTPPWT	399
90	QASQDISNYLN	241	AASTLRS	331	QQSYQTPLT	400
91	QASQDISNYLN	241	AASTLQS	299	QQSYTTPPT	401
92	QASQDISNYLN	241	AASSLHS	332	QQSYSTPQT	371
93	RASQGIRNDLN	291	AASNLQS	330	QQANSFPIT	402
94	RASQGINTWLA	292	AASSLQS	294	QQSYSTPYT	344
95	QASQDISNYLN	241	AASTLQS	299	QQSYTVPPT	403
96	QASQDIRYFLN	757	AASTLQS	299	QQQDSFPLT	404

Table 10. HCDRs using the IMGT Numbering Scheme

CD8 Binder	H-CDR1		H-CDR2		H-CDR3	
	Sequence	SEQ ID NO:	Sequence	SEQ ID NO:	Sequence	SEQ ID NO:
1	GGTFSSYA	758	IDPSDGNT	817	AKERAAAGYYYYMDV	878
2	GGTFNTYA	759	IDPSSGGT	818	AKEHAAGTYYYYMDV	879
3	GGTFSSYA	758	IDPSSGGT	819	AKERAAAGYYYYMDV	878
4	GYTFTGY	760	INPNNGDT	820	AKEGYYYYGMDV	880
5	GYTFTDY	761	INPNSSGT	821	AKEGDYYYGMDA	881
6	GYTFTRYD	762	INPNDGST	822	ARERGGMPDY	882
7	GYTFTSYA	763	INPNSSGT	821	ARGHGIPKY	883
8	GYTFTSY	764	MNPNSGNT	823	ARVRSGSPQH	884
9	GHTFSRHY	765	MNPNSGNT	823	ARGGPWIVDAFDI	885
10	GYTFTSYG	766	ISAHNGVT	824	ARGIAVAGTDY	886
11	GGTFSNTD	767	INPSGGST	825	AREATWGPYYYYMDV	887
12	GYTFTRSY	768	ISPYNGNT	826	VRNKDGLQH	888
13	GDTFTGY	769	INPNSSGT	827	AKDAKRVGYYYYMDV	889
14	GYTFTRY	770	INPNSSGT	821	ARLVGGSPDY	890
15	GYTFTNYD	771	INPNSSGT	821	ARGAMVDY	891
16	GGTFSNTD	767	INPSDGD	828	ARGNYVGSYYYGMDV	892
17	GYTFTNYY	772	INPNSSGT	827	ARDSRGDWYFDL	893
18	GYGFTSYS	773	IDPSSGGST	829	TRHGGRGLADY	894
19	GYTFTSRD	774	IDPKSGDT	830	ARLKELSSILDAFDI	895
20	GYTFTSYD	775	INPGAGSS	831	ARERFGTGYYYYMDV	896
21	GFTFSNSD	776	ISGDGGTT	832	ARVIGEMVDDAFDL	897
22	GYTFTGY	760	INPNSSGT	827	ARERLFGTYYYYMDV	898
23	GYTFTTYD	777	IPIFGTA	833	ARADGELTDY	899
24	GFTFSSYT	778	I-GTGGGI	834	ARHHLPAHYYYYYMDV	900
25	GGTFSRYD	779	INPNSSGT	827	ARDVPAGRYYYYYMDV	901
26	GNTFTSY	780	INPSDGST	835	AKDRGVGRYYYYMDV	902
27	GGTFSRYA	781	INPSDGST	835	AKDSRYGRYYYYMDV	903
28	GGTFSNYA	782	INPNSSGSP	836	AKEIVVGPYYYYMDV	904
29	GGTFTRYA	783	INPNSSGT	827	ARGMVRGPYYYYMDV	905
30	GGTFSSYA	758	INPSGGST	825	AREGVTGPYYYYMDV	906
31	GGTFSRFD	784	INPSDGST	835	ARDAAAGTRYYYYYMDV	907

32	GGTFSSHA	785	INPSGGST	825	ARELYSSTYYYYMDV	908
33	GGTFSSYA	758	INPNTGGT	837	ARALYSGPYYYYYMDV	909
34	GFTFSNSD	776	ISGSGGST	838	AKEHAAGTYYYYMDV	879
35	GGTFGSYG	786	ISGYNGDT	839	ARDSLVGRYYYYMDV	910
36	GYIFTDYD	787	ISADNGNT	840	ARRSELDY	911
37	GYTFTSYH	788	ISPNSGAT	841	ARGDDNDY	912
38	GYTFTSYD	775	INPNSGNT	842	ARGEEVDY	913
39	GYTFTSYP	789	INPSGGST	825	ARGRRVPDY	914
40	GYTFTDYY	761	INPKSGAT	843	ARGKVTTDY	915
41	GFTFSSFE	790	ISESGDSS	844	ASGRELIEY	916
42	GFTFDDYA	791	I-GTGGGT	845	ARVYDFPDV	917
43	GYTFTDSY	792	MNPSNGDT	846	ARSTYSHIDY	918
44	GYTFTNYY	772	ISPSDGST	847	AREDSSGFDY	919
45	GYTFMNY	793	INPSGGST	825	ARDQGGGFDY	920
46	GYTFTSYY	764	FDPEDGET	848	ARDQGWGMDV	921
47	GYTFTSYY	764	INPKSGRT	849	ARLTEGIPDY	922
48	GYTLNDYY	794	INPGGGST	850	ARDRYGPFDY	923
49	GYTFTSYD	775	MNPKTGDT	851	TRLVAGGAPDY	924
50	GYTFTGY	760	IDPSDGYT	852	ARDGFTGDIAY	925
51	GYTFTGY	760	INPNSGGT	821	ARVDDSSSPDY	926
52	GYTFTGY	760	IMPISGTT	853	TTGPDGTEVDY	927
52	GYTFTNHY	795	MNPNSGNT	823	ASSESGSDLDY	928
54	GYTFTNYY	772	MSPTSGDT	854	AREVEIEGYMDV	929
55	GYTFTSYY	764	INPNSGDT	827	AKDLDDDWYMDV	930
56	GYTFTSYY	764	IDPSGDIT	855	TTDSTTWDAFDI	931
57	GYTFTDYY	761	INPNSGGT	821	ARVLVGSGSPDY	932
58	GYTFTENE	796	IETSGGST	856	AREAAAGLDFQH	933
59	GYTFASYD	797	INPNSGGT	821	ARANSWDAMVIDY	934
60	GFTFSNSD	776	ISGSGVTT	857	AREHSSSWYTFDY	935
61	GYTFTAYY	798	INPNSGGT	821	ARDDDSSGYLDY	936
62	GYTFTNYY	772	INPSGGST	825	ARASGDYMDLIDY	937
63	GYTFTDYH	799	INPDSSGT	858	ALVGSSGYLAPTH	938
64	GYPFTDYY	800	MNPNSGNT	823	ARVRGDGYNLGDY	939
65	GYTFSDYY	801	INPNSGGT	821	ARDVDTAMGAGDY	940
66	GYTFTDYY	761	INPSGSA	859	ARVARWGYGDYPDY	941
67	GDTFTTHD	802	ISPSDGST	847	ARDRNGDYGGMDV	942
68	GDTFTNYY	803	INPISGGT	860	AREGLSSWYVLDY	943
69	GYTFTSYD	775	ISADNGDT	861	ARDGSHYGYGMDV	944
70	GYTFTSYD	775	ISPIYGTP	862	ASPGPEGYYYGMDV	945
71	GYTFTDNY	804	MNPNSGNT	823	ASYHWDYGDYRFDY	946
72	GYTFTSYY	764	MNPNSGNT	823	ARVEIDYGDSPPDY	947
73	GGTSSSYA	805	INPSDGD	828	ARGAEWELRYAFDI	948
74	GYTFTTYD	777	INPSGGTT	863	ARETYYGLYYYGMDV	949
75	GYTFTSYD	775	MNPKSGNT	864	ARAPSLRGYSYGPDY	950
76	GGTFTSYD	806	INPSGGST	825	AKDRQERYYYYYMDV	951
77	GYTFTSYD	775	INPSDGST	835	AKDRSYGDYYYGMDV	952
78	GGTFTSYD	806	INPGGGNA	865	AREVFSENYYYYMDV	953
79	GYTFTSYY	764	INPSDGST	835	AREWDYTHYYYGMDV	954
80	GNTFTSHW	807	FDPEDGET	848	ARGDSSGYQQYFDY	955
81	GYTFTSYD	775	ITPVFGIA	866	ARGSWDSSSWYIPEY	956
82	GFTFSDYD	808	INPRGGST	867	ASLVWGGAYYYYMDV	957
83	GYTFTSYG	766	MNPNNGDT	868	TTPVFSGSYWYFDP	958
84	GYTFTSYD	775	INPSGGGT	869	TTDQAVAGPYYYGMDV	959

85	GGTFSSYA	758	INPGSGNT	870	ARDRWLAGPYYYGMDV	960
86	GYMFTGHD	809	IPIFGTP	871	ARVMGPVDYYYYGMDV	961
87	GYIFSNYD	810	INPSDGST	835	ARDLGPFGSYYYYMDV	962
88	GFTFSSYA	811	INGDGDDT	872	AREGVVPPYYYYMDV	963
89	GYTFTTYY	812	IDPNSGDT	873	ARSSGWSRYYYYYMDV	964
90	GSTFTNYQ	813	INPSGGST	825	ARDNGMTTGYYYGMDV	965
91	GYTFTSYD	775	INPSGGST	825	ARDRAMVTGYYYGMDV	966
92	GYTFTSYD	775	VNPSDGNT	874	ARDRGYGDRGYYYGMDV	967
93	GGTLSSYD	814	INTYNGNT	875	ATSPKATADYYYYMDV	968
94	GYTFTSYD	775	INPSDGIT	876	TTSTVTPSYYYGMDV	969
95	GYTFTSHA	815	INPRDGT	877	AREPVAGTGYYYGMDV	970
96	GGTFNSYG	816	MNPNSGNT	823	ARDNLAGFWSYYYYGMDV	971
97	GRTFSGYV	1066	ISRGGLST	1067	AARDSDLYEITAASNIDS	1068

Table 11. LCDRs using the IMGT Numbering Scheme

CD8 Binder	L-CDR1		L-CDR2		L-CDR3	
	Sequence	SEQ ID NO:	Sequence	SEQ ID NO:	Sequence	SEQ ID NO:
1	QSISSY	972	AAS	1021	QQSYSTPLT	333
2	QDISNY	973	AAS	1021	QQSYSNLVS	334
3	QSISSY	972	AAS	1021	QQSYSTPLT	333
4	QSLHNSNGYNY	974	LGS	1022	MQALQTPFT	335
5	QSLHNSNGYNY	974	LGS	1022	MQGLQTPHT	336
6	QSIERN	975	KAS	1023	QQTYSAPL	337
7	QSLHNSNGYNY	974	LGS	1022	MQLQTPLT	338
8	QSVSASD	976	GAS	1024	QQYGDSPGS	339
9	QDIGNY	977	AAS	1021	QQANSFPPT	340
10	QSIETH	978	GAS	1024	QQYGNST	341
11	QTISNY	979	AAS	1021	QQSYSTPPT	342
12	QGIRND	980	DAS	1025	QQSYSSPYT	343
13	QSIERN	981	AAS	1021	QQSYSTPYT	344
14	QSLHNSNGYNY	974	LGS	1022	MQGAHWPPPT	345
15	QGISDS	982	GAS	1024	QQSYRTPYT	346
16	QSIERN	981	AAS	1021	QESFTTQWT	347
17	QDIHNY	983	DAS	1025	QQANSFPPT	340
18	QDISNY	973	SAS	1026	QQRSNWPLYT	348
19	QSIDDW	984	AAS	1021	QQAISFPIT	349
20	QDISNY	973	SAS	1026	QQSYSSPFT	350
21	QSIETW	985	AAS	1021	QQAISFPLT	351
22	QSIERN	981	AAS	1021	QQSYTFPIT	352
23	QSLHNSNGYNY	974	DAS	1025	QQYYSYPPT	353
24	QDISNY	973	AAS	1021	QQSYSAPLT	354
25	QDISNY	973	AAS	1021	QQSFSTFYT	355
26	QDISNY	973	AAS	1021	QQSYSIPFT	356
27	QSINRF	986	AAS	1021	QQSYSTPYT	344
28	QSISSY	972	AAS	1021	QQSYSTPLT	333
29	QDISNY	973	AAS	1021	QQSYSTPIT	357
30	QSVSTY	987	AAS	1021	QQSYTIPST	358
31	QDIKAY	988	AAS	1021	QQSYSAPPT	359
32	QGITNY	989	GAS	1024	QQSYSTPWT	360
33	QSISSY	972	AAS	1021	QQSYSTPLT	333

34	QDIHNY	983	AAS	1021	QQSYTTPLT	361
35	QDISNY	973	SAF	1027	QQSYSAPIT	362
36	QSISSY	972	SAS	1026	QQRSNWPPVT	363
37	QDISNF	990	DAS	1025	QQSYSIPIT	364
38	QGISNN	991	EAS	1028	QQSYSTPLT	333
39	QSLHNSNGYNY	974	GAS	1024	MQGLQPPGT	365
40	QSIIRS	992	AAS	1021	QQSYNHFRT	366
41	QDISNY	973	DAS	1025	QRSDSTPLT	367
42	HDISKS	993	GAS	1024	QQLNSYPRT	368
43	QDIGAY	994	AAS	1021	QQSYSIPYT	369
44	QSISSY	972	AAS	1021	QQSYSTPYT	344
45	QGIRSY	995	GAS	1024	QQSYSTPYT	344
46	QSISSY	972	AAS	1021	QQTYSTPYT	370
47	QNIQW	996	AAS	1021	QQSYSTPQT	371
48	QTISYY	997	AAS	1021	QQSYRTPYT	346
49	QSLHNSNGYNY	974	MGS	1029	MQGTHWPT	372
50	QNIINNY	998	GAS	1024	QQTFLPYT	373
51	QTISTY	999	DAS	1025	QQSYSTPYT	344
52	RGIGND	1000	DAS	1025	QQGYNMPLT	374
52	QTIGNY	1001	GAS	1024	QQTYSAPLT	375
54	QFIGSW	1002	AAS	1021	QQSYSPWT	376
55	QSISSW	1003	DAS	1025	QQTYSTPYI	377
56	QGISNN	991	DAS	1025	QQSYSSPWT	378
57	QSVLYSSNNKNY	1004	WAS	1030	QQYASAPRT	379
58	QSISSY	972	KTS	1031	QQSFTIPYT	380
59	QGISSY	1005	GAS	1024	QQSYSTPLT	333
60	QSIDDW	984	DAS	1025	QQSYSTPLT	333
61	QGISNY	1006	SAS	1026	QQTyrTPPT	381
62	QSIRNY	1007	SAS	1026	QQSYSTPLT	333
63	QNIIRLY	1008	AAS	1021	QQLTTPFT	382
64	QDIRKF	1009	AAS	1021	QQLNGYPT	383
65	QSISSY	972	TAS	1032	QQSYSLPLT	384
66	QDISNY	973	DAS	1025	QQTYYTPRT	385
67	QNVRSW	1010	AAS	1021	QQSYNTPYT	386
68	QGIGND	1011	AAS	1021	QQSYAPPPT	387
69	QSIISNW	1012	GAS	1024	QQSYSTPPT	342
70	QSLHNSNGYNY	974	LGS	1022	MQGLQTPLT	388
71	QSISSY	972	LAS	1033	QQSDSIPVT	389
72	QDISNY	973	STS	1034	QQSYSTPYN	390
73	ESIGSW	1013	AAS	1021	QQSYSTPYT	344
74	QSIISNY	981	AAS	1021	QQSYSTPLT	333
75	QSVTSNY	1014	GAS	1024	QHYGSSPA	391
76	QSISSY	972	AAS	1021	QQSYSTPLT	333
77	QGISSY	1005	AAS	1021	QQSYSTPPT	342
78	QDIGNY	977	AAS	1021	QQAYTYPYT	392
79	QDISNY	973	GAS	1024	QQSYTTPNT	393
80	QGISNY	1006	AAS	1021	QQSYSTPYT	344
81	QGISNG	1015	DAS	1025	QQSYSTPFT	394
82	QNIIRNY	1016	GAS	1024	QQSYSTPLT	333
83	LDINNY	1017	KAS	1023	QQSYSMLPT	395
84	QDISNY	973	AAS	1021	QQSYTTPWT	396
85	QDISNY	973	AAS	1021	QQSYSSPLT	397
86	QDISNY	973	KAS	1023	QQSYSDPLT	398

87	QDISNY	973	GAS	1024	QQSYSAPIT	362
88	QSISNY	981	AAS	1021	QQSYTTPLT	361
89	QNIGNY	1018	AAS	1021	QQSYSTPPWT	399
90	QDISNY	973	AAS	1021	QQSYQTPLT	400
91	QDISNY	973	AAS	1021	QQSYTTPPT	401
92	QDISNY	973	AAS	1021	QQSYSTPQT	371
93	QGIRND	980	AAS	1021	QQANSFPIT	402
94	QGINTW	1019	AAS	1021	QQSYSTPYT	344
95	QDISNY	973	AAS	1021	QQSYTVPPT	403
96	QDIRYF	1020	AAS	1021	QQDDSFPLT	404

Table 12

CD8 Binder Variant	Full Binder Sequence	SEQ ID NO:
98	QVQLVQSGAEVKKPGASVKVSCKASGGTFSSYAISWVRQAPGQGLEWMGIIDPSDGNTNYAQN FQGRVTMTRDTSTSTVYMESSLRSEDAVYYCAKERAAGYNYMDVWGQGTITVSSGGGGSG GGGGGGGGSDIQMTQSPSSLSASVGDRVTITCRASQSISSYLNWYQQKPGKAPKLLIYAASLQSGV PSRFSGSGSGTDFTLTISSLQPEDFATYYCQQSYSTPLTFGGGKVEIKR	1035
99	DIQMTQSPSSLSASVGDRVTITCRASQSISSYLNWYQQKPGKAPKLLIYAASLQSGVPSRFSGSGSG TDFTLTISSLQPEDFATYYCQQSYSTPLTFGGGKVEIKRGGGGSGGGGGSGGGGSQVQLVQSGAEV KKPGASVKVSCKASGGTFSSYAISWVRQAPGQGLEWMGIIDPSDGNTNYAQNFGQGRVTMTRDTS TSTVYMESSLRSEDAVYYCAKERAAGYNYMDVWGQGTITVSS	1036
100	QVQLVQSGAEVKKPGASVKVSCKASGGTFSSYAISWVRQAPGQGLEWMGIIDPSDGNTNYAQN FQGRVTMTRDTSTSTVYMESSLRSEDAVYYCAKERAAGYNYMDVWGQGTITVSSGGGGSG GGGGGGGGSDIQMTQSPSSLSASVGDRVTITCRASQSISSYLNWYQQKPGKAPKLLIYAASLQSGV PSRFSGSGSGTDFTLTISSLQPEDFATYYCQQSYSTPLTFGGGKVEIKR	1037
101	DIQMTQSPSSLSASVGDRVTITCRASQSISSYLNWYQQKPGKAPKLLIYAASLQSGVPSRFSGSGSG TDFTLTISSLQPEDFATYYCQQSYSTPLTFGGGKVEIKRGGGGSGGGGGSGGGGSQVQLVQSGAEV KKPGASVKVSCKASGGTFSSYAISWVRQAPGQGLEWMGIIDPSDGNTNYAQNFGQGRVTMTRDTS TSTVYMESSLRSEDAVYYCAKERAAGYNYMDVWGQGTITVSS	1038
102	QVQLVQSGAEVKKPGASVKVSCKASGGTFSSYAISWVRQAPGQGLEWMGIIDPSDGNTNYAQN FQGRVTMTRDTSTSTVYMESSLRSEDAVYYCAKERAAGYNYMDVWGQGTITVSSGGGGSG GGGGGGGGSDIQMTQSPSSLSASVGDRVTITCRASQSISSYLNWYQQKPGKAPKLLIYAASLQSGV PSRFSGSGSGTDFTLTISSLQPEDFATYYCQQSYSTPLTFGGGKVEIKR	1037
103	DIQMTQSPSSLSASVGDRVTITCRASQSISSYLNWYQQKPGKAPKLLIYAASLQSGVPSRFSGSGSG TDFTLTISSLQPEDFATYYCQQSYSTPLTFGGGKVEIKRGGGGSGGGGGSGGGGSQVQLVQSGAEV KKPGASVKVSCKASGGTFSSYAISWVRQAPGQGLEWMGIIDPSDGNTNYAQNFGQGRVTMTRDTS TSTVYMESSLRSEDAVYYCAKERAAGYNYMDVWGQGTITVSS	1038
104	DIVMTQSPSLPVPTEGEPASISCRSSQSLHSGNYLDWYLQKPGQSPQLLIYLSNRRASGVPDRFS GSGSGTDFTLKISRVEAEDVGVYYCMQGLQTPHTFGGQGTVEIKRGGGGSGGGGGSGGGGSQVQL VQSGAEVKKPGASVKVSCKASGYTFDYYIQWVRQAPGQGLEWMGWINPNSGGTSAQKFQGR VTMTRDTSTSTVYMESSLRSEDAVYYCAKEGDYNYGMDAWGQGTMTVSS	1039
105	QVQLVQSGAEVKKPGASVKVSCKASGYTFDYYIQWVRQAPGQGLEWMGWINPNSGGTSAQK FQGRVTMTRDTSTSTVYMESSLRSEDAVYYCAKEGDYNYGMDAWGQGTMTVSSGGGGSGG GGGGGGSDIVMTQSPSLPVPTEGEPASISCRSSQSLHSGNYLDWYLQKPGQSPQLLIYLSNR ASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYCMQGLQTPHTFGGQGTVEIKR	1040
106	DIVMTQSPSLPVPTEGEPASISCRSSQSLHSGNYLDWYLQKPGQSPQLLIYLSNRRASGVPDRFS GSGSGTDFTLKISRVEAEDVGVYYCMQGLQTPHTFGGQGTVEIKRGGGGSGGGGGSGGGGSQVQL VQSGAEVKKPGASVKVSCKASGYTFDYYIQWVRQAPGQGLEWMGWINPNSGGTSAQKFQGR VTMTRDTSTSTVYMESSLRSEDAVYYCAKEGDYNYGMDAWGQGTMTVSS	1039

107	QVQLVQSGAEVKKPGASVKVSCKASGYTFTDYIQQWVRQAPGQCLEWMGWINPNSGGTSYAQK FQGRVTMTRDTSTSTVYMESSLRSED AVYYCAKEGDYYYGMDAWGQGTMTVSSGGGGSGG GGSGGGGSDIVMTQSPSLPVTGPGEASISCRSSQSLHNSNGYNYLDWYLQKPGQSPQLLIYLGSR ASGVDPDRFSGSGSGTDFTLKISRVEAEDVGVYYCMQGLQTPHTFGCGTKVEIKR	1041
108	DIVMTQSPSLPVTGPGEASISCRSSQSLHNSNGYNYLDWYLQKPGQSPQLLIYLGSRASGVDPDRF SGSGSGTDFTLKISRVEAEDVGVYYCMQGLQTPHTFGCGTKVEIKRGGGGSGGGSGGGGSQVQL VQSGAEVKKPGASVKVSCKASGYTFTDYIQQWVRQAPGQCLEWMGWINPNSGGTSYAQKFQGR VTMTRDTSTSTVYMESSLRSED AVYYCAKEGDYYYGMDAWGQGTMTVSS	1042
109	QVQLVQSGAEVKKPGASVKVSCKASGYTFTDYIQQWVRQAPGQCLEWMGWINPNSGGTSYAQK FQGRVTMTRDTSTSTVYMESSLRSED AVYYCAKEGDYYYGMDAWGQGTMTVSSGGGGSGG GGSGGGGSDIVMTQSPSLPVTGPGEASISCRSSQSLHNSNGYNYLDWYLQKPGQSPQLLIYLGSR ASGVDPDRFSGSGSGTDFTLKISRVEAEDVGVYYCMQGLQTPHTFGCGTKVEIKR	1041
110	DIVMTQSPSLPVTGPGEASISCRSSQSLHNSNGYNYLDWYLQKPGQSPQLLIYLGSRASGVDPDRF SGSGSGTDFTLKISRVEAEDVGVYYCMQGLQTPHTFGCGTKVEIKRGGGGSGGGSGGGGSQVQL VQSGAEVKKPGASVKVSCKASGYTFTDYIQQWVRQAPGQCLEWMGWINPNSGGTSYAQKFQGR VTMTRDTSTSTVYMESSLRSED AVYYCAKEGDYYYGMDAWGQGTMTVSS	1042
111	QVQLVQSGAEVKKPGASVKVSCKASGYTFTSYMHVVRQAPGQCLEWMGGFDPEDGETIYAQK FQGRVTMTRDTSTSTVYMESSLRSED AVYYCARDQGWGMDVWGQGTITVSSGGGGSGGG GGSGGGSDIQMTQSPSSLSASVGDRVTITCRASQSISSYLNWYQQKPGKAPKLLIYAASLSQGVPS RFSGSGSGTDFTLTISLQPEDFATYYCQQTSTPYTFGCGTKLEIKR	1043
112	QVQLVQSGAEVKKPGASVKVSCKASGYTFTSYMHVVRQAPGQCLEWMGGFDPEDGETIYAQK FQGRVTMTRDTSTSTVYMESSLRSED AVYYCARDQGWGMDVWGQGTITVSSGGGGSGGG GGSGGGSDIQMTQSPSSLSASVGDRVTITCRASQSISSYLNWYQQKPGKAPKLLIYAASLSQGVPS RFSGSGSGTDFTLTISLQPEDFATYYCQQTSTPYTFGCGTKLEIKR	1043
113	QVQLVQSGAEVKKPGASVKVSCKASGYTFTSYMHVVRQAPGQCLEWMGGFDPEDGETIYAQK FQGRVTMTRDTSTSTVYMESSLRSED AVYYCARDQGWGMDVWGQGTITVSSGGGGSGGG GGSGGGSDIQMTQSPSSLSASVGDRVTITCRASQSISSYLNWYQQKPGKAPKLLIYAASLSQGVPS RFSGSGSGTDFTLTISLQPEDFATYYCQQTSTPYTFGCGTKLEIKR	1044
114	DIQMTQSPSSLSASVGDRVTITCRASQSISSYLNWYQQKPGKAPKLLIYAASLSQGVPSRFSGSGS TDFTLTISLQPEDFATYYCQQTSTPYTFGCGTKLEIKRGGGGSGGGSGGGGSQVQLVQSGAEV KKPGASVKVSCKASGYTFTSYMHVVRQAPGQCLEWMGGFDPEDGETIYAQKFQGRVTMTRDT STSTVYMESSLRSED AVYYCARDQGWGMDVWGQGTITVSS	1045
115	DIQMTQSPSSLSASVGDRVTITCRASQSISSYLNWYQQKPGKAPKLLIYAASLSQGVPSRFSGSGS TDFTLTISLQPEDFATYYCQQTSTPYTFGCGTKLEIKRGGGGSGGGSGGGGSQVQLVQSGAEV KKPGASVKVSCKASGYTFTSYMHVVRQAPGQCLEWMGGFDPEDGETIYAQKFQGRVTMTRDT STSTVYMESSLRSED AVYYCARDQGWGMDVWGQGTITVSS	1045
116	DIQMTQSPSSLSASVGDRVTITCRASQSISSYLNWYQQKPGKAPKLLIYAASLSQGVPSRFSGSGS TDFTLTISLQPEDFATYYCQQTSTPYTFGCGTKLEIKRGGGGSGGGSGGGGSQVQLVQSGAEV KKPGASVKVSCKASGYTFTSYMHVVRQAPGQCLEWMGGFDPEDGETIYAQKFQGRVTMTRDT STSTVYMESSLRSED AVYYCARDQGWGMDVWGQGTITVSS	1046
117	DIQMTQSPSSLSASVGDRVTITCRASQTIGNYVNWYQQKPGKAPKLLIYGASNLHTGVPSRFSGSGS GTDFTLTISLQPEDFATYYCQQTYSAPLTFGGGKVEIKRGGGGSGGGSGGGGSQVQLVQSGAE VKKPGASVKVSCKASGYTFTNHYMHVVRQAPGQCLEWMGWMNPNSGNTGYAQKFQGRVTM TRDTSTSTVYMESSLRSED AVYYCASSESGDLDYWGQGLTVTVSS	1047
118	QVQLVQSGAEVKKPGASVKVSCKASGYTFTNHYMHVVRQAPGQCLEWMGWMNPNSGNTGY AQKFQGRVTMTRDTSTSTVYMESSLRSED AVYYCASSESGDLDYWGQGLTVTVSSGGGGSGG GGSGGGSDIQMTQSPSSLSASVGDRVTITCRASQTIGNYVNWYQQKPGKAPKLLIYGASNLHTG VPSRFSGSGSGTDFTLTISLQPEDFATYYCQQTYSAPLTFGGGKVEIKR	1048
119	DIQMTQSPSSLSASVGDRVTITCRASQTIGNYVNWYQQKPGKAPKLLIYGASNLHTGVPSRFSGSGS GTDFTLTISLQPEDFATYYCQQTYSAPLTFGGGKVEIKRGGGGSGGGSGGGGSQVQLVQSGAE VKKPGASVKVSCKASGYTFTNHYMHVVRQAPGQCLEWMGWMNPNSGNTGYAQKFQGRVTM TRDTSTSTVYMESSLRSED AVYYCASSESGDLDYWGQGLTVTVSS	1047
120	QVQLVQSGAEVKKPGASVKVSCKASGYTFTNHYMHVVRQAPGQCLEWMGWMNPNSGNTGY AQKFQGRVTMTRDTSTSTVYMESSLRSED AVYYCASSESGDLDYWGQGLTVTVSSGGGGSGG	1049

	GGSGGGSDIQMTQSPSSLSASVGDRVTITCRASQTIGNYVNWYQQKPGKAPKLLIYGASNLHTG VPSRFSGSGSGDFTLTISLQPEDFATYYCQQTYSAPLTFGCGTKVEIKR	
121	DIQMTQSPSSLSASVGDRVTITCRASQTIGNYVNWYQQKPGKAPKLLIYGASNLHTGVPSRFSGSGS GTDFTLTISLQPEDFATYYCQQTYSAPLTFGCGTKVEIKRGGGGSGGGSGGGGSQVQLVQSGAE VKKPGASVKVSCASGYFTFNHYMHVVRQAPGQCLEWMGWMNPNPNSGNTGYAQKFQGRVTM TRDTSTSTVYMESSLRSEDATVYYCASSESGDLDYWGQGLTVTVSS	1050
122	QVQLVQSGAEVKKPGASVKVSCASGYFTFNHYMHVVRQAPGQCLEWMGWMNPNPNSGNTGY AQKFQGRVTMTRDTSTSTVYMESSLRSEDATVYYCASSESGDLDYWGQGLTVTVSSGGGGSGG GGSGGGSDIQMTQSPSSLSASVGDRVTITCRASQTIGNYVNWYQQKPGKAPKLLIYGASNLHTG VPSRFSGSGSGDFTLTISLQPEDFATYYCQQTYSAPLTFGCGTKVEIKR	1049
123	DIQMTQSPSSLSASVGDRVTITCRASQTIGNYVNWYQQKPGKAPKLLIYGASNLHTGVPSRFSGSGS GTDFTLTISLQPEDFATYYCQQTYSAPLTFGCGTKVEIKRGGGGSGGGSGGGGSQVQLVQSGAE VKKPGASVKVSCASGYFTFNHYMHVVRQAPGQCLEWMGWMNPNPNSGNTGYAQKFQGRVTM TRDTSTSTVYMESSLRSEDATVYYCASSESGDLDYWGQGLTVTVSS	1050
124	QVQLVQSGAEVKKPGASVKVSCASGYFTFNHYMHVVRQAPGQCLEWMGIINPNPNSGNTGYAQ KFQGRVTMTRDTSTSTVYMESSLRSEDATVYYCASSESGDLDYWGQGLTVTVSSGGGGSGGGG SGGGSDIQMTQSPSSLSASVGDRVTITCRASQTIGNYVNWYQQKPGKAPKLLIYGASNLHTGVPS RFSGSGSGDFTLTISLQPEDFATYYCQQTYSAPLTFGGGKVEIKR	1051
125	DIQMTQSPSSLSASVGDRVTITCRASQTIGNYVNWYQQKPGKAPKLLIYGASNLHTGVPSRFSGSGS GTDFTLTISLQPEDFATYYCQQTYSAPLTFGGGKVEIKRGGGGSGGGSGGGGSQVQLVQSGAE VKKPGASVKVSCASGYFTFNHYMHVVRQAPGQCLEWMGIINPNPNSGNTGYAQKFQGRVTMTR DTSTSTVYMESSLRSEDATVYYCASSESGDLDYWGQGLTVTVSS	1052
126	QVQLVQSGAEVKKPGASVKVSCASGYFTFNHYMHVVRQAPGQCLEWMGIINPNPNSGNTGYAQ KFQGRVTMTRDTSTSTVYMESSLRSEDATVYYCASSESGDLDYWGQGLTVTVSSGGGGSGGGG SGGGSDIQMTQSPSSLSASVGDRVTITCRASQTIGNYVNWYQQKPGKAPKLLIYGASNLHTGVPS RFSGSGSGDFTLTISLQPEDFATYYCQQTYSAPLTFGGGKVEIKR	1051
127	DIQMTQSPSSLSASVGDRVTITCRASQTIGNYVNWYQQKPGKAPKLLIYGASNLHTGVPSRFSGSGS GTDFTLTISLQPEDFATYYCQQTYSAPLTFGGGKVEIKRGGGGSGGGSGGGGSQVQLVQSGAE VKKPGASVKVSCASGYFTFNHYMHVVRQAPGQCLEWMGIINPNPNSGNTGYAQKFQGRVTMTR DTSTSTVYMESSLRSEDATVYYCASSESGDLDYWGQGLTVTVSS	1052
128	QVQLVESGGGLVQAGGSLRLSCAASGRTFSGYVMGWFRQAPGKQRKFVAISRGLSTSYADSVK GRFTISRDNKNTVFLQMNTLKPEDATVYYCAADRSLDYEITAASNIDSWGQGLTVTVSS	1060

Table 13

CD8 Fusion Proteins	Full Fusion Protein Sequence	SEQ ID NO:
NivG.002 _1	MKKINEGLLDKILSAFNTVIALGSIIVMNIMIIQNYTRSTDNQAVIKDALQGIQQQIKGLADKIG TEIGPKVSLIDTSSTITIPANIGLLGSKISQSTASINENVNEKCKFTLPPLKIECNISCPNPLPFREYRPQ TEGVSNLVGLPNNICLKQTSNQILKPKLISYTLPPVVGQSGTCITDPLLAMDEGYFAYSHLERIGCSR GVSKQRIIGVGEVLDRGDEVPSLFMTNVWTPPNPNTVYHCSAVYNNEFYVLCVAVSTVGPILNS TYWSGSLMMTRLAVKPKSNGGGYNQHQLALRSIEKGRYDKVMPYGPSGIKQGDITLYPAVGFLV RTEFKYNDNCPITKCYSKPENCRSLMGIIRPNSHYILRSGLLYNLSGDNENPKVVFIEISDQRLSIGS PSKIYDSLQGPVFYQASFSWDTMIKFGDVLTVNPLVVNWRNNTVISRPGQSQCPRFNTCPAICAE GVYNDAFLIDRINWISAGVFLDSNATAANPVFTVKDNEILYRAQLASEDTNAQKTITNCFLLKNKI WCISLVEIYDTGDNVIRPKLFAVKIPEQCTGGGGSGGGSGGGGSQVQLVQSGAEVKKPGASVKV SCKASGTFSSYAIWVRQAPGQGLEWMGIIDPSDGNNTYANFNQGRVTMTRDTSTSTVYME SSLRSEDATVYYCAKERAAGYYYMDVWGQGTITVTVSSGGGGSGGGSGGGSDIQMTQSPS LSASVGDRVTITCRASQSISSYLNWYQQKPGKAPKLLIYAASSLQSGVPSRFSGSGSGDFTLTISL QPEDFATYYCQSYSTPLTFGGGKVEIKR	1053
NivG.002 _5	MKKINEGLLDKILSAFNTVIALGSIIVMNIMIIQNYTRSTDNQAVIKDALQGIQQQIKGLADKIG TEIGPKVSLIDTSSTITIPANIGLLGSKISQSTASINENVNEKCKFTLPPLKIECNISCPNPLPFREYRPQ	1054

	TEGVSNLVGLPNNICLQKTSNQILKPKLISYTLPVVGQSGTCITDPLLAMDEGYFAYSHLERIGSCSR GVSKQRIIGVGEVLDRGDEVPSLFMTNVWTPPNPNTVYHCSAVYNNEFYVLCVAVSTVGDPIILNS TYWSGSLMMTRLAVKPKSNGGGYNQHQLALRSIEKGRYDKVMPYGPSGIKQGDITLYFPAVGFLV RTEFKYNDNSNCPITKCQYSKPENCRLSMGIKPNSHYILRSGLLKYNLSDGENPKVVFIEISDQRLSIG PSKIYDSLQGPVQYQASFSWDTMIKFGDVLTVNPLVNVWRNNTVISRPGQSQCPRFNTCPAICAE GVYNDAFLIDRINWISAGVFLDSNATAANPVFTVFKDNEILYRAQLASEDTNAQKTIITNCFLLKNKI WCISLVEIYDGTGDNVIRPKLFAVKIPEQCTGGGGSGGGSGGGGSQVQLVQSGAEVKKPGASVKV SCKASGYTFTDYIQWVRQAPGQGLEWMGWINPNSGGTSAQKFQGRVTMTTRDTSTSTVYME LSSLRSEDATVYYCAKEGDYGYMDAWGQGTMTVTVSSGGGGSGGGSGGGGSQVQLVQSGAEVKKPGASVKV PVTGPGEASISCRSSQSLHSGNYLDWYQLKPGQSPQLLIYLSNRASGVPDRFSGSGSGTDFTL KISRVEAEDVGVYCYMQGLQTPHTFGQGTKEIKR	
NivG.002 _46	MKKINEGLLDSKILSAFNTVIALLSGSIIVMNIMIIQNYTRSTDNQAVIKDALQGIQQQIKGLADKIG TEIGPKVSLIDTSSITIPANIGLLGSKISQSTASINENVNEKCKFTLPPLKIECNISCPNPLPFREYRPQ TEGVSNLVGLPNNICLQKTSNQILKPKLISYTLPVVGQSGTCITDPLLAMDEGYFAYSHLERIGSCSR GVSKQRIIGVGEVLDRGDEVPSLFMTNVWTPPNPNTVYHCSAVYNNEFYVLCVAVSTVGDPIILNS TYWSGSLMMTRLAVKPKSNGGGYNQHQLALRSIEKGRYDKVMPYGPSGIKQGDITLYFPAVGFLV RTEFKYNDNSNCPITKCQYSKPENCRLSMGIKPNSHYILRSGLLKYNLSDGENPKVVFIEISDQRLSIG PSKIYDSLQGPVQYQASFSWDTMIKFGDVLTVNPLVNVWRNNTVISRPGQSQCPRFNTCPAICAE GVYNDAFLIDRINWISAGVFLDSNATAANPVFTVFKDNEILYRAQLASEDTNAQKTIITNCFLLKNKI WCISLVEIYDGTGDNVIRPKLFAVKIPEQCTGGGGSGGGSGGGGSQVQLVQSGAEVKKPGASVKV SCKASGYTFTSYMHVVRQAPGQGLEWMGGFDPEDGETIYAQKFQGRVTMTTRDTSTSTVYME LSSLRSEDATVYYCARDQGWGMDVWGQGTTVTVSSGGGGSGGGSGGGGSQVQLVQSGAEVKKPGASVKV ASVGDRVTITCRASQSSSYLNWYQQKPGKAPKLLIYAASSLQSGVPSRFSGSGSGTDFTLTISLQ EDFATYYCQQTSTPYTFGQGTKEIKR	1055
NivG.002 _52	MKKINEGLLDSKILSAFNTVIALLSGSIIVMNIMIIQNYTRSTDNQAVIKDALQGIQQQIKGLADKIG TEIGPKVSLIDTSSITIPANIGLLGSKISQSTASINENVNEKCKFTLPPLKIECNISCPNPLPFREYRPQ TEGVSNLVGLPNNICLQKTSNQILKPKLISYTLPVVGQSGTCITDPLLAMDEGYFAYSHLERIGSCSR GVSKQRIIGVGEVLDRGDEVPSLFMTNVWTPPNPNTVYHCSAVYNNEFYVLCVAVSTVGDPIILNS TYWSGSLMMTRLAVKPKSNGGGYNQHQLALRSIEKGRYDKVMPYGPSGIKQGDITLYFPAVGFLV RTEFKYNDNSNCPITKCQYSKPENCRLSMGIKPNSHYILRSGLLKYNLSDGENPKVVFIEISDQRLSIG PSKIYDSLQGPVQYQASFSWDTMIKFGDVLTVNPLVNVWRNNTVISRPGQSQCPRFNTCPAICAE GVYNDAFLIDRINWISAGVFLDSNATAANPVFTVFKDNEILYRAQLASEDTNAQKTIITNCFLLKNKI WCISLVEIYDGTGDNVIRPKLFAVKIPEQCTGGGGSGGGSGGGGSQVQLVQSGAEVKKPGASVKV SCKASGYTFTNHYMHVVRQAPGQGLEWMGWMNPNNSGNTGYAQKFQGRVTMTTRDTSTSTVY MELSSLRSEDATVYYCASSESGDLDYWGQGTTLTVSSGGGGSGGGSGGGGSQVQLVQSGAEVKKPGASVKV SASVGDRVTITCRASQTIGNYVNWYQQKPGKAPKLLIYGASNLHTGVPSRFSGSGSGTDFTLTISL QPEDFATYYCQQTYSAPLTFGGGTKEIKR	1056
NivG.002 _97	MKKINEGLLDSKILSAFNTVIALLSGSIIVMNIMIIQNYTRSTDNQAVIKDALQGIQQQIKGLADKIG TEIGPKVSLIDTSSITIPANIGLLGSKISQSTASINENVNEKCKFTLPPLKIECNISCPNPLPFREYRPQ TEGVSNLVGLPNNICLQKTSNQILKPKLISYTLPVVGQSGTCITDPLLAMDEGYFAYSHLERIGSCSR GVSKQRIIGVGEVLDRGDEVPSLFMTNVWTPPNPNTVYHCSAVYNNEFYVLCVAVSTVGDPIILNS TYWSGSLMMTRLAVKPKSNGGGYNQHQLALRSIEKGRYDKVMPYGPSGIKQGDITLYFPAVGFLV RTEFKYNDNSNCPITKCQYSKPENCRLSMGIKPNSHYILRSGLLKYNLSDGENPKVVFIEISDQRLSIG PSKIYDSLQGPVQYQASFSWDTMIKFGDVLTVNPLVNVWRNNTVISRPGQSQCPRFNTCPAICAE GVYNDAFLIDRINWISAGVFLDSNATAANPVFTVFKDNEILYRAQLASEDTNAQKTIITNCFLLKNKI WCISLVEIYDGTGDNVIRPKLFAVKIPEQCTGGGGSGGGSGGGGSQVQLVESGGGLVQAGGSLRL SCAASGRTFSGYVMGWFRQAPGKQKQFVAISRGLSTYADSVKGRFTISRDNKNTVFLQMN TLKPEDATVYYCAADRSDLYEITAASNIDSWGQGTTLTVSS	1069
NivG.002 _1_2	KKINEGLLDSKILSAFNTVIALLSGSIIVMNIMIIQNYTRSTDNQAVIKDALQGIQQQIKGLADKIGTE IGPKVSLIDTSSITIPANIGLLGSKISQSTASINENVNEKCKFTLPPLKIECNISCPNPLPFREYRPQTE GVSNLVGLPNNICLQKTSNQILKPKLISYTLPVVGQSGTCITDPLLAMDEGYFAYSHLERIGSCSRGV SKQRIIGVGEVLDRGDEVPSLFMTNVWTPPNPNTVYHCSAVYNNEFYVLCVAVSTVGDPIILNSTY WSGSLMMTRLAVKPKSNGGGYNQHQLALRSIEKGRYDKVMPYGPSGIKQGDITLYFPAVGFLVRT EFKYNDNSNCPITKCQYSKPENCRLSMGIKPNSHYILRSGLLKYNLSDGENPKVVFIEISDQRLSIGSP KIYDSLQGPVQYQASFSWDTMIKFGDVLTVNPLVNVWRNNTVISRPGQSQCPRFNTCPAICAEV	1070

	YNDAFLIDRINWISAGVFLDSNATAANPVFTVFKDNEILYRAQLASEDTNAQKTITNCFLLKNIWC ISLVEIYDTGDNVIRPKLFAVKIPEQCTGGGGSGGGSGGGGSQVQLVQSGAEVKKPGASVKVSC KASGGTFSSYAIWVRQAPGQGLEWMGIIDPSDGNNTYANFQGRVTMTTRDTSTSTVYMELSSL RSEDVAVYYCAKERAAGYYYMDVWVGQGTTVTVSSGGGGSGGGSGGGGSDIQMTQSPSSLS ASVGDRVTITCRASQSISSYLNWYQQKPGKAPKLLIYAASSLQSGVPSRFSGSGSGDTFTLTISLQPE EDFATYYCQSYSTPLTFGGGKVEIKR	
NivG.002 _5_2	KKINEGLLDSKILSAFNTVIALLSIVIMNIMIIQNYTRSTDNQAVIKDALQGIQQIKGLADKIGTE IGPKVSLIDTSSTITIPANIGLLGSKISQSTASINENVNEKCKFTLPPLKHECNISCPNPLPFREYRPQTE GVSNLVGLPNNICLQKTSNQILPKLISYTLPVVGQSGTCITDPLLAMDEGYFAYSHLERIGSCSRGV SKQRIIGVGEVLDRGDEVPSLFMTNVWTPPNPNTVYHCSAVYNNEFYVLCVSTVGDPIINSTY WSGSLMMTRLAVKPKSNGGGYNQHLALRSIEKGRYDKVMPIYGPYSGIKQGDITLYFPAVGFLVRT EFKYNSDNCPIKQYKSPENCRLSMGIRPNSHYILRSGLLKYNLSDGENPKVVFIEISDQRLSIGSPS KIYDSLQGPVIFYQASFSWDTMIKFGDVLTVNPLVWNRNNTVISRPGQSQCPRENTCPAICAEV YNDAFLIDRINWISAGVFLDSNATAANPVFTVFKDNEILYRAQLASEDTNAQKTITNCFLLKNIWC ISLVEIYDTGDNVIRPKLFAVKIPEQCTGGGGSGGGSGGGGSQVQLVQSGAEVKKPGASVKVSC KASGYTFTDYYIQWVRQAPGQGLEWMGWINPNSGGTSAQKFQGRVTMTTRDTSTSTVYMELS SLRSEDVAVYYCAKEGDYGYMDAWGQGTMTVTVSSGGGGSGGGSGGGGSDIVMTQSPSLP VTPGEPASISCRSSQSLHSNGYNLDWYLQKPGQSPQLLIYLSNRASGVPDRFSGSGSGDTFTLK ISRVEAEDVGYYCMQGLQTPHFGQGTKEIKR	1071
NivG.002 _46_2	KKINEGLLDSKILSAFNTVIALLSIVIMNIMIIQNYTRSTDNQAVIKDALQGIQQIKGLADKIGTE IGPKVSLIDTSSTITIPANIGLLGSKISQSTASINENVNEKCKFTLPPLKHECNISCPNPLPFREYRPQTE GVSNLVGLPNNICLQKTSNQILPKLISYTLPVVGQSGTCITDPLLAMDEGYFAYSHLERIGSCSRGV SKQRIIGVGEVLDRGDEVPSLFMTNVWTPPNPNTVYHCSAVYNNEFYVLCVSTVGDPIINSTY WSGSLMMTRLAVKPKSNGGGYNQHLALRSIEKGRYDKVMPIYGPYSGIKQGDITLYFPAVGFLVRT EFKYNSDNCPIKQYKSPENCRLSMGIRPNSHYILRSGLLKYNLSDGENPKVVFIEISDQRLSIGSPS KIYDSLQGPVIFYQASFSWDTMIKFGDVLTVNPLVWNRNNTVISRPGQSQCPRENTCPAICAEV YNDAFLIDRINWISAGVFLDSNATAANPVFTVFKDNEILYRAQLASEDTNAQKTITNCFLLKNIWC ISLVEIYDTGDNVIRPKLFAVKIPEQCTGGGGSGGGSGGGGSQVQLVQSGAEVKKPGASVKVSC KASGYTFTSYMHVVRQAPGQGLEWMGGFDPEDGETIYAQKFQGRVTMTTRDTSTSTVYMELS SLRSEDVAVYYCARDQGWGMDVWVGQGTTVTVSSGGGGSGGGSGGGGSDIQMTQSPSSLSA SVGDRVTITCRASQSISSYLNWYQQKPGKAPKLLIYAASSLQSGVPSRFSGSGSGDTFTLTISLQPE DFATYYCQQTSTPYTFGQGTKEIKR	1072
NivG.002 _52_2	KKINEGLLDSKILSAFNTVIALLSIVIMNIMIIQNYTRSTDNQAVIKDALQGIQQIKGLADKIGTE IGPKVSLIDTSSTITIPANIGLLGSKISQSTASINENVNEKCKFTLPPLKHECNISCPNPLPFREYRPQTE GVSNLVGLPNNICLQKTSNQILPKLISYTLPVVGQSGTCITDPLLAMDEGYFAYSHLERIGSCSRGV SKQRIIGVGEVLDRGDEVPSLFMTNVWTPPNPNTVYHCSAVYNNEFYVLCVSTVGDPIINSTY WSGSLMMTRLAVKPKSNGGGYNQHLALRSIEKGRYDKVMPIYGPYSGIKQGDITLYFPAVGFLVRT EFKYNSDNCPIKQYKSPENCRLSMGIRPNSHYILRSGLLKYNLSDGENPKVVFIEISDQRLSIGSPS KIYDSLQGPVIFYQASFSWDTMIKFGDVLTVNPLVWNRNNTVISRPGQSQCPRENTCPAICAEV YNDAFLIDRINWISAGVFLDSNATAANPVFTVFKDNEILYRAQLASEDTNAQKTITNCFLLKNIWC ISLVEIYDTGDNVIRPKLFAVKIPEQCTGGGGSGGGSGGGGSQVQLVQSGAEVKKPGASVKVSC KASGYTFTNHYMHVVRQAPGQGLEWMGMNPNNSGNTGYAQKFQGRVTMTTRDTSTSTVYM ELSSLRSEDVAVYYCASSESGDLDYWGQGTTLTVTVSSGGGGSGGGSGGGGSDIQMTQSPSSLSA SVGDRVTITCRASQTIGNYVWYQQKPGKAPKLLIYGASNLHTGVPSRFSGSGSGDTFTLTISLQPE EDFATYYCQQTYSAPLTFGGGKVEIKR	1073
NivG.002 _97_2	KKINEGLLDSKILSAFNTVIALLSIVIMNIMIIQNYTRSTDNQAVIKDALQGIQQIKGLADKIGTE IGPKVSLIDTSSTITIPANIGLLGSKISQSTASINENVNEKCKFTLPPLKHECNISCPNPLPFREYRPQTE GVSNLVGLPNNICLQKTSNQILPKLISYTLPVVGQSGTCITDPLLAMDEGYFAYSHLERIGSCSRGV SKQRIIGVGEVLDRGDEVPSLFMTNVWTPPNPNTVYHCSAVYNNEFYVLCVSTVGDPIINSTY WSGSLMMTRLAVKPKSNGGGYNQHLALRSIEKGRYDKVMPIYGPYSGIKQGDITLYFPAVGFLVRT EFKYNSDNCPIKQYKSPENCRLSMGIRPNSHYILRSGLLKYNLSDGENPKVVFIEISDQRLSIGSPS KIYDSLQGPVIFYQASFSWDTMIKFGDVLTVNPLVWNRNNTVISRPGQSQCPRENTCPAICAEV YNDAFLIDRINWISAGVFLDSNATAANPVFTVFKDNEILYRAQLASEDTNAQKTITNCFLLKNIWC ISLVEIYDTGDNVIRPKLFAVKIPEQCTGGGGSGGGSGGGGSQVQLVESGGGLVQAGGSLRLSC	1074

	AASGRTFSGYVMGWFRQAPGKQRKFVAAISRGGGLSTSYADSVKGRFTISRDNKNTVFLQMNTL KPEDTAVYYCAADRSDLYEITAASNIDSWGQGT LVTSS	
--	---	--

CLAIMS:

1. An antibody or antigen binding fragment thereof that specifically binds CD8 α or CD8 β , comprising a heavy chain variable region and/or a light chain variable region, wherein the heavy chain variable region comprises three heavy chain complementarity determining regions (HCDR1, HCDR2, and HCDR3), and the light chain variable region comprises three light chain complementarity determining regions (LCDR1, LCDR2, and LCDR3), wherein the HCDR1, HCDR2, HCDR3, LCDR1, LCDR2, and LCDR3, respectively, comprise:
 - a) SEQ ID NOs: 1, 70, 148, 240, 294, and 333, respectively;
 - b) SEQ ID NOs: 2, 71, 149, 241, 294, and 334, respectively;
 - c) SEQ ID NOs: 3, 72, 148, 240, 294, and 333, respectively;
 - d) SEQ ID NOs: 4, 73, 649, 242, 295, and 335, respectively;
 - e) SEQ ID NOs: 5, 74, 150, 242, 295, and 336, respectively;
 - f) SEQ ID NOs: 6, 75, 151, 243, 296, and 337, respectively;
 - g) SEQ ID NOs: 7, 76, 152, 242, 295, and 338, respectively;
 - h) SEQ ID NOs: 8, 77, 153, 244, 297, and 339, respectively;
 - i) SEQ ID NOs: 9, 77, 154, 245, 298, and 340, respectively;
 - j) SEQ ID NOs: 10, 78, 650, 246, 297, and 341, respectively;
 - k) SEQ ID NOs: 11, 79, 155, 247, 299, and 342, respectively;
 - l) SEQ ID NOs: 12, 80, 156, 248, 300, and 343, respectively;
 - m) SEQ ID NOs: 13, 81, 157, 249, 294, and 344, respectively;
 - n) SEQ ID NOs: 14, 76, 158, 242, 295, and 345, respectively;
 - o) SEQ ID NOs: 15, 76, 159, 250, 301, and 346, respectively;
 - p) SEQ ID NOs: 11, 82, 160, 249, 294, and 347, respectively;
 - q) SEQ ID NOs: 16, 83, 161, 251, 302, and 340, respectively;
 - r) SEQ ID NOs: 17, 84, 162, 241, 303, and 348, respectively;
 - s) SEQ ID NOs: 18, 85, 163, 252, 304, and 349, respectively;
 - t) SEQ ID NOs: 19, 86, 164, 241, 305, and 350, respectively;
 - u) SEQ ID NOs: 20, 87, 165, 253, 299, and 351, respectively;
 - v) SEQ ID NOs: 4, 88, 166, 249, 299, and 352, respectively;
 - w) SEQ ID NOs: 21, 89, 167, 242, 306, and 353, respectively;

- x) SEQ ID NOs: 22, 90, 168, 241, 307, and 354, respectively;
- y) SEQ ID NOs: 23, 91, 169, 241, 299, and 355, respectively;
- z) SEQ ID NOs: 24, 92, 170, 241, 299, and 356, respectively;
- aa) SEQ ID NOs: 25, 93, 171, 254, 308, and 344, respectively;
- bb) SEQ ID NOs: 26, 94, 172, 240, 294, and 333, respectively;
- cc) SEQ ID NOs: 617, 91, 173, 241, 299, and 357, respectively;
- dd) SEQ ID NOs: 1, 79, 174, 255, 294, and 358, respectively;
- ee) SEQ ID NOs: 28, 95, 175, 256, 294, and 359, respectively;
- ff) SEQ ID NOs: 29, 79, 176, 257, 309, and 360, respectively;
- gg) SEQ ID NOs: 1, 96, 177, 240, 294, and 333, respectively;
- hh) SEQ ID NOs: 20, 97, 149, 251, 299, and 361, respectively;
- ii) SEQ ID NOs: 30, 98, 178, 241, 310, and 362, respectively;
- jj) SEQ ID NOs: 31, 99, 179, 240, 311, and 363, respectively;
- kk) SEQ ID NOs: 32, 100, 180, 258, 312, and 364, respectively;
- ll) SEQ ID NOs: 19, 101, 181, 259, 313, and 333, respectively;
- mm) SEQ ID NOs: 33, 102, 182, 242, 314, and 365, respectively;
- nn) SEQ ID NOs: 34, 103, 183, 260, 315, and 366, respectively;
- oo) SEQ ID NOs: 35, 104, 184, 241, 302, and 367, respectively;
- pp) SEQ ID NOs: 36, 105, 185, 261, 316, and 368, respectively;
- qq) SEQ ID NOs: 37, 106, 186, 262, 294, and 369, respectively;
- rr) SEQ ID NOs: 38, 107, 187, 263, 294, and 344, respectively;
- ss) SEQ ID NOs: 39, 108, 188, 264, 317, and 344, respectively;
- tt) SEQ ID NOs: 40, 109, 189, 240, 294, and 370, respectively;
- uu) SEQ ID NOs: 8, 110, 190, 265, 299, and 371, respectively;
- vv) SEQ ID NOs: 41, 111, 191, 266, 299, and 346, respectively;
- ww) SEQ ID NOs: 19, 112, 192, 242, 318, and 372, respectively;
- xx) SEQ ID NOs: 4, 113, 193, 267, 309, and 373, respectively;
- yy) SEQ ID NOs: 4, 114, 194, 268, 302, 344, and 453, respectively;
- zz) SEQ ID NOs: 42, 115, 195, 269, 319, 374, and 454, respectively;
- aaa) SEQ ID NOs: 43, 77, 196, 270, 320, and 375, respectively;
- bbb) SEQ ID NOs: 44, 116, 197, 271, 299, and 376, respectively;
- ccc) SEQ ID NOs: 40, 117, 198, 272, 302, and 377, respectively;
- ddd) SEQ ID NOs: 40, 118, 199, 259, 302, and 378, respectively;

eee) SEQ ID NOs: 45, 114, 200, 273, 321, and 379, respectively;
fff) SEQ ID NOs: 46, 119, 201, 240, 322, and 380, respectively;
ggg) SEQ ID NOs: 47, 120, 202, 274, 309, and 333, respectively;
hhh) SEQ ID NOs: 48, 121, 203, 252, 302, and 333, respectively;
iii) SEQ ID NOs: 49, 122, 204, 275, 311, and 381, respectively;
jjj) SEQ ID NOs: 44, 123, 205, 276, 311, and 333, respectively;
kkk) SEQ ID NOs: 50, 124, 206, 277, 299, and 382, respectively;
lll) SEQ ID NOs: 51, 77, 207, 278, 294, and 383, respectively;
mmm) SEQ ID NOs: 52, 125, 208, 240, 323, and 384, respectively;
nnn) SEQ ID NOs: 34, 126, 209, 279, 324, and 385, respectively;
ooo) SEQ ID NOs: 53, 127, 210, 280, 294, and 386, respectively;
ppp) SEQ ID NOs: 54, 128, 211, 281, 294, and 387, respectively;
qqq) SEQ ID NOs: 19, 129, 212, 282, 317, and 342, respectively;
rrr) SEQ ID NOs: 19, 130, 213, 242, 295, and 388, respectively;
sss) SEQ ID NOs: 55, 77, 214, 240, 325, and 389, respectively;
ttt) SEQ ID NOs: 8, 77, 215, 241, 326, and 390, respectively;
uuu) SEQ ID NOs: 56, 131, 216, 283, 294, and 344, respectively;
vvv) SEQ ID NOs: 57, 132, 217, 249, 327, and 333, respectively;
www) SEQ ID NOs: 19, 133, 218, 284, 297, and 391, respectively;
xxx) SEQ ID NOs: 58, 79, 219, 240, 294, and 333, respectively;
yyy) SEQ ID NOs: 19, 95, 220, 285, 299, and 342, respectively;
zzz) SEQ ID NOs: 58, 134, 221, 286, 294, and 392, respectively;
aaaa) SEQ ID NOs: 40, 93, 222, 241, 309, and 393, respectively;
bbbb) SEQ ID NOs: 59, 135, 223, 275, 299, and 344, respectively;
cccc) SEQ ID NOs: 19, 136, 224, 287, 302, and 394, respectively;
dddd) SEQ ID NOs: 60, 137, 225, 288, 309, and 333, respectively;
eeee) SEQ ID NOs: 10, 138, 226, 289, 328, and 395, respectively;
ffff) SEQ ID NOs: 19, 139, 227, 241, 329, and 396, respectively;
gggg) SEQ ID NOs: 1, 140, 228, 241, 294, and 397, respectively;
hhhh) SEQ ID NOs: 61, 141, 229, 241, 328, and 395, respectively;
iiii) SEQ ID NOs: 62, 93, 230, 241, 316, and 362, respectively;
jjjj) SEQ ID NOs: 63, 142, 231, 249, 330, 361, and 490, respectively;
kkkk) SEQ ID NOs: 64, 143, 232, 290, 299, and 399, respectively;

- llll) SEQ ID NOs: 65, 79, 233, 241, 331, and 400, respectively;
mmmm) SEQ ID NOs: 66, 79, 234, 241, 299, and 401, respectively;
nnnn) SEQ ID NOs: 19, 144, 235, 241, 332, and 371, respectively;
oooo) SEQ ID NOs: 67, 145, 236, 291, 330, and 402, respectively;
pppp) SEQ ID NOs: 19, 146, 237, 292, 294, and 344, respectively;
qqqq) SEQ ID NOs: 68, 147, 238, 241, 299, and 403, respectively;
rrrr) SEQ ID NOs: 69, 77, 239, 293, 299, and 404, respectively; and
ssss) SEQ ID NOs: 1061, 1062, and 1063, respectively.
2. The antibody or antigen binding fragment thereof of claim 1, comprising a heavy chain variable region (VH) having an amino acid sequence with at least 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to a sequence selected from SEQ ID NOs: 405-498 or 1060.
 3. The antibody or antigen binding fragment thereof of claim 1 or 2, comprising a light chain variable region (VL) having an amino acid sequence with at least 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to a sequence selected from SEQ ID NOs: 499-591.
 4. The antibody or antigen binding fragment thereof of claim 2 or 3, comprising a VH having an amino acid sequence with at least 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to a sequence selected from SEQ ID NOs: 405-498 and a VL having an amino acid sequence with at least 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to a sequence selected from SEQ ID NOs: 499-591.
 5. The antibody or antigen binding fragment thereof of any one of claims 2-4, wherein the VH is selected from SEQ ID NOs: 405, 408, 448, 455, and 1060, and the VL is selected from 499, 503, 542, and 549.
 6. The antibody or antigen binding fragment thereof of claim 1, wherein the HCDR1, HCDR2, HCDR3, LCDR1, LCDR2, and LCDR3, respectively, comprise:
 - a) SEQ ID NOs: 1, 70, 148, 240, 294, 333, respectively;
 - b) SEQ ID NOs: 5, 74, 150, 242, 295, 336, respectively;
 - c) SEQ ID NOs: 40, 109, 189, 240, 294, 370, respectively;
 - d) SEQ ID NOs: 43, 77, 196, 270, 320, 375, respectively;

7. The antibody or antigen binding fragment thereof of any one of claims 1-6, which binds to human CD8 α or CD8 β .
8. The antibody or antigen binding fragment thereof of any one of claims 1-7, which binds to a human CD8 α homodimer composed of two α chains.
9. The antibody or antigen binding fragment thereof of any one of claims 1-7, which binds to a human CD8 heterodimer composed of one α chain and one β chain.
10. The antibody or antigen binding fragment thereof of any one of claims 1-9, wherein the antigen binding fragment is a Fab, Fab', F(ab')₂, Fd, scFv, (scFv)₂, scFv-Fc, sdAb, VHH, or Fv fragment.
11. The antibody or antigen binding fragment thereof of claim 10, wherein the antigen binding fragment is a scFv.
12. The antibody or antigen binding fragment thereof of claim 11, wherein the VH is on the N-terminal side of the LH.
13. The antibody or antigen binding fragment thereof of claim 11, wherein the VL is on the N-terminal side of the VH.
14. The antibody or antigen binding fragment thereof of any one of claims 11-13, wherein the scFv comprises a linker connecting the VH and VL.
15. The antibody or antigen binding fragment thereof of claim 14, wherein the linker comprises the amino acid sequence set forth in SEQ ID NOs: 625-627 and 645-648.
16. The antibody or antigen binding fragment thereof of any one of claims 1-15, which binds to human CD8 α or human CD8 β with a K_D of about 400 nM or lower.
17. The antibody or antigen binding fragment thereof of any one of claims 1-15, which binds to M. nemestrina CD8 α or M. nemestrina CD8 β with a K_D of about 5 nM to about 500 nM .
18. The antibody or antigen binding fragment thereof of any one of claims 1-15, which binds to human CD8 α or CD8 β with a K_D of about 5 nM to about 500 nM.
19. An isolated polynucleotide encoding the antibody or antigen binding fragment thereof of any of claims 1-18.
20. An isolated vector comprising the polynucleotide of claim 19.

21. An isolated host cell comprising the polynucleotide of claim 19, and/or the vector of claim 20.
22. A fusion protein comprising a Henipavirus glycoprotein G (G protein) or a biologically active portion thereof and at least one antibody or antigen binding fragment thereof of any one of claims 1-18, wherein the antibody or antigen binding fragment is fused to the C-terminus of the G protein or the biologically active portion thereof.
23. The fusion protein of claim 22, wherein the antibody or antigen binding fragment thereof is fused to the G protein via a peptide linker.
24. The fusion protein of claim 23, wherein the peptide linker comprises a polypeptide that is up to 65 amino acids in length.
25. The fusion protein of any one of claims 22-24, wherein the peptide linker comprises from about 2 to 65 amino acids.
26. The fusion protein of any one of claims 22-25, wherein the peptide linker comprises a polypeptide that is 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64 or 65 amino acids in length.
27. The fusion protein of any one of claims 22-26, wherein the peptide linker comprises GS, GGS, GGGS (SEQ ID NO: 645), GGGGS (SEQ ID NO: 626 and 627), GGGGGS (SEQ ID NO: 625), or combinations thereof.
28. The fusion protein of any one of claims 22-27, wherein the peptide linker comprises (GGS)_n, wherein n is 1 to 10.
29. The fusion protein of any one of claims 22-27, wherein the peptide linker comprises (GGS)_n, wherein n is 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10.
30. The fusion protein of any one of claims 22-27, wherein the peptide linker comprises (GGGS)_n (SEQ ID NO: 645), wherein n is 1 to 10.
31. The protein of any one of claims 22-27, wherein the peptide linker comprises (GGGS)_n (SEQ ID NO: 645), wherein n is 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10.
32. The fusion protein of any one of claims 22-27, wherein the peptide linker comprises (GGGGS)_n (SEQ ID NO: 626), wherein n is 1 to 10.

33. The fusion protein of any one of claims 22-27, wherein the peptide linker comprises (GGGGS)_n (SEQ ID NO: 626), wherein n is 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10.
34. The fusion protein of any one of claims 22-27, wherein the peptide linker comprises (GGGGGS)_n (SEQ ID NO: 27), wherein n is 1 to 6.
35. The fusion protein of any one of claims 22-27, wherein the peptide linker comprises (GGGGGS)_n (SEQ ID NO: 27), wherein n is 1, 2, 3, 4, 5, or 6.
36. The fusion protein of any one of claims 22-35, wherein the antigen binding fragment is a scFv.
37. The fusion protein of any one of claims 22-36, wherein the G protein or a biologically active portion thereof is a wild-type Nipah virus G glycoprotein (NiV-G) or a functionally active variant or a biologically active portion thereof.
38. The fusion protein of any one of claims 22-37, wherein the NiV-G variant or biologically active portion thereof comprises an amino acid sequence having at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to SEQ ID NO: 600, SEQ ID NO: 618, or SEQ ID NO: 628.
39. The fusion protein of any one of claims 22-38, wherein the NiV-G protein is a biologically active portion that is truncated and lacks up to 40 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein SEQ ID NO: 600, SEQ ID NO: 618, or SEQ ID NO: 628.
40. The fusion protein of any one of claims 22-37, wherein the NiV-G protein is a biologically active portion that is truncated at the N-terminus of wild-type NiV-G and has the sequence set forth in any of SEQ ID NOS: 601-606 or 629-634, or an amino acid sequence having at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to SEQ ID NOS: 601-606 or 629-634.

41. The fusion protein of any one of claims 22-39, wherein the NiV-G protein has a 5 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO: 600, SEQ ID NO: 618, or SEQ ID NO: 628).
42. The fusion protein of any one of claims 22-39, wherein the NiV-G protein has a 10 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO: 600, SEQ ID NO: 618, or SEQ ID NO: 628).
43. The fusion protein of any one of claims 22-39, wherein the NiV-G protein has a 15 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO: 600, SEQ ID NO: 618, or SEQ ID NO: 628).
44. The fusion protein of any one of claims 22-39, wherein the NiV-G protein has a 20 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO: 600, SEQ ID NO: 618, or SEQ ID NO: 628).
45. The fusion protein of any one of claims 22-39, wherein the NiV-G protein has a 25 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO: 600, SEQ ID NO: 618, or SEQ ID NO: 628).
46. The fusion protein of any one of claims 22-39, wherein the NiV-G protein has a 30 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO: 600, SEQ ID NO: 618, or SEQ ID NO: 628).
47. The fusion protein of any one of claims 22-39, wherein the NiV-G protein has a 34 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO: 600, SEQ ID NO: 618, or SEQ ID NO: 628).
48. The fusion protein of any one of claims 22-47, wherein the NiV-G variant or biologically active portion thereof exhibits reduced binding to Ephrin B2 or Ephrin B3.
49. The fusion protein of any one of claims 22-39, wherein the NiV-G variant or biologically active portion thereof comprises one or more amino acid substitutions corresponding to amino acid substitutions selected from E501A, W504A, Q530A and E533A with reference to the numbering set forth in SEQ ID NO: 618.
50. The fusion protein any one of claims 22-39, wherein the NiV-G variant or biologically active portion thereof comprises the amino acid substitutions corresponding to amino acid substitutions selected from E501A, W504A,

Q530A and E533A with reference to the numbering set forth in SEQ ID NO: 618.

51. The fusion protein of any one of claims 22-39, wherein the NiV-G variant or biologically active portion thereof comprises an amino acid sequence having at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to SEQ ID NO: 607.
52. The fusion protein of any one of claims 22-39, wherein the NiV-G variant or biologically active portion thereof comprises an amino acid sequence having at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to SEQ ID NO: 635.
53. The fusion protein of any one of claims 22-39, wherein the NiV-G protein has a 34 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO: 600, SEQ ID NO: 618, or SEQ ID NO: 628) and one or more amino acid substitutions corresponding to amino acid substitutions selected from E501A, W504A, Q530A and E533A with reference to the numbering set forth in SEQ ID NO: 618.
54. The fusion protein of any one of claims 22-54, in which the protein is pseudotyped onto a lentiviral particle.
55. The fusion protein of any one of claims 22-55, wherein the antibody or antigen binding fragment thereof binds to cell surface molecule present on a CD8⁺ T cell.
56. A viral vector comprising:
 - i) a henipavirus F protein molecule or biologically active portion thereof;
 - ii) a henipavirus envelope glycoprotein G (G protein) or a biologically active portion thereof; and
 - iii) at least one antibody or antigen binding fragment thereof of any one of claims 1-18,

wherein the antibody or antigen binding fragment thereof is attached to the C-terminus of the G protein or the biologically active portion thereof.

57. The viral vector of claim 56, wherein the antibody or antigen binding fragment thereof binds a cell surface molecule present on a target cell.
58. The viral vector of claim 57, wherein the target cell is a CD8+ T cell.
59. The viral vector of any one of claims 56-58, wherein the antibody or antigen binding fragment thereof binds an antigen or portion thereof present on a target cell.
60. The viral vector of any one of claims 56-59, wherein the antigen binding fragment is a scFv.
61. The viral vector of any one of claims 56-59, wherein the F protein or the biologically active portion thereof is a wild-type Nipah virus F (NiV-F) protein or a functionally active variant or a biologically active portion thereof.
62. The viral vector of any one of claims 56-59, wherein the NiV-F-protein or the functionally active variant or biologically active portion thereof comprises the amino acid sequence set forth in SEQ ID NO: 593, or an amino acid sequence having at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to SEQ ID NO: 593.
63. The viral vector of any one of claims 56-62, wherein the NiV-F protein is a biologically active portion thereof that has a 20 amino acid truncation at or near the C-terminus of the wild-type NiV-F protein (SEQ ID NO: 593).
64. The viral vector of any one of claims 56-62, wherein the NiV-F protein has an amino acid sequence having at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to SEQ ID NO: 596.
65. The viral vector of any one of claims 56-62, wherein the NiV-F protein is a biologically active portion thereof that comprises:

- i) a 20 amino acid truncation at or near the C-terminus of the wild-type NiV-F protein (SEQ ID NO: 593); and
 - ii) a point mutation on an N-linked glycosylation site.
66. The viral vector of any one of claims 56-62, wherein the NiV-F protein is a biologically active portion thereof that has a 22 amino acid truncation at or near the C-terminus of the wild-type NiV-F protein (SEQ ID NO: 593).
67. The viral vector of claim 66, wherein the NiV-F protein has an amino acid sequence that is encoded by a sequence of nucleotides encoding a sequence having at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to SEQ ID NO: 599.
68. The viral vector of claim 57, wherein the NiV-F protein has an amino acid sequence having at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to SEQ ID NO: 613.
69. The viral vector of any one of claims 56-68, wherein the F-protein or the biologically active portion thereof comprises an F1 subunit or a fusogenic portion thereof.
70. The viral vector of claim 69, wherein the F1 subunit is a proteolytically cleaved portion of the F0 precursor.
71. The viral vector of claim 70, wherein the F1 subunit comprises the sequence set forth in SEQ ID NO: 595, or an amino acid sequence having at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 9%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to SEQ ID NO: 595.
72. The viral vector of any one of claims 56-71, wherein the G protein or a biologically active portion thereof is a wild-type Nipah virus G glycoprotein (NiV-G) or a functionally active variant or a biologically active portion thereof.

73. The viral vector of any one of claims 56-72, wherein the NiV-G variant or biologically active portion thereof comprises an amino acid sequence having at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to SEQ ID NO: 600, SEQ ID NO: 618, or SEQ ID NO: 628.
74. The viral vector of any one of claims 56-73, wherein the NiV-G protein is a biologically active portion that is truncated and lacks up to 40 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO: 600, SEQ ID NO: 618, or SEQ ID NO: 628).
75. The viral vector of any one of claims 56-72, wherein the NiV-G protein is a biologically active portion that is truncated at the N-terminus of wild-type NiV-G and has the sequence set forth in any of SEQ ID NOS: 601-606 or 629-634, or an amino acid sequence having at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to SEQ ID NOS: 601-606 or 629-634.
76. The viral vector of any one of claims 56-73, wherein the NiV-G protein has a 5 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO: 600, SEQ ID NO: 618, or SEQ ID NO: 628).
77. The viral vector of any one of claims 56-73, wherein the NiV-G protein has a 10 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO: 600, SEQ ID NO: 618, or SEQ ID NO: 628).
78. The viral vector of any one of claims 56-73, wherein the NiV-G protein has a 15 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO: 600, SEQ ID NO: 618, or SEQ ID NO: 628).
79. The viral vector of any one of claims 56-73, wherein the NiV-G protein has a 20 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO: 600, SEQ ID NO: 618, or SEQ ID NO: 628).

80. The viral vector of any one of claims 56-73, wherein the NiV-G protein has a 25 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO: 600, SEQ ID NO: 618, or SEQ ID NO: 628).
81. The viral vector of any one of claims 56-73, wherein the NiV-G protein has a 30 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO: 600, SEQ ID NO: 618, or SEQ ID NO: 628).
82. The viral vector of any one of claims 56-73, wherein the NiV-G protein has a 34 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO: 600, SEQ ID NO: 618, or SEQ ID NO: 628).
83. The viral vector of any one of claims 56-82, wherein the NiV-G variant or biologically active portion thereof exhibits reduced binding to Ephrin B2 or Ephrin B3.
84. The viral vector of any one of claims 56-83, wherein the NiV-G variant or biologically active portion thereof comprises one or more amino acid substitutions corresponding to amino acid substitutions selected from E501A, W504A, Q530A and E533A with reference to the numbering set forth in SEQ ID NO: 618.
85. The viral vector any one of claims 56-84, wherein the NiV-G variant or biologically active portion thereof comprises the amino acid substitutions corresponding to amino acid substitutions selected from E501A, W504A, Q530A and E533A with reference to the numbering set forth in SEQ ID NO: 618.
86. The viral vector of any one of claims 56-85, wherein the NiV-G variant or biologically active portion thereof comprises an amino acid sequence having at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to SEQ ID NO: 607.
87. The viral vector of any one of claims 56-73, wherein the NiV-G variant or biologically active portion thereof comprises an amino acid sequence having at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at

least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to SEQ ID NO: 635.

88. The viral vector of any one of claims 56-87, wherein the NiV-G protein has a 34 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO: 600, SEQ ID NO: 618, or SEQ ID NO: 628) and one or more amino acid substitutions corresponding to amino acid substitutions selected from E501A, W504A, Q530A and E533A with reference to the numbering set forth in SEQ ID NO: 618.
89. The viral vector of claim 88, wherein the NiV-F protein is a biologically active portion thereof that has a 22 amino acid truncation at or near the C-terminus of the wild-type NiV-F protein (SEQ ID NO: 593).
90. The viral vector of any one of claims 56-89, wherein the NiV-G variant or biologically active portion thereof exhibits reduced binding to Ephrin B2 or Ephrin B3.
91. The viral vector of any one of claims 56-90, wherein the viral vector is a lentiviral vector.
92. The viral vector of any one of claims 56-91, wherein the viral vector is a retroviral vector.
93. The viral vector of any one of claims 56-92, wherein the viral vector comprises one or more viral components other than the F protein molecule and the G protein.
94. The viral vector of claim 93, wherein the one or more viral components comprise a viral protein selected from one or more of Gag, Pol, Rev, and Tat.
95. The viral vector of claim 93 or 94, wherein the one or more viral components comprises one or more endogenous nucleic acid sequences encoding a chimeric antigen receptor (CAR).
96. The viral vector of claim 93 or 94, wherein the one or more viral components comprises one or more endogenous nucleic acid sequences encoding a T cell receptor (TCR).
97. The viral vector of any of claims 94-96, wherein the one or more viral components comprises one or more of the following nucleic acid sequences: a 5' LTR (e.g., comprising U5 and lacking a functional U3 domain), a Psi

packaging element (Psi), a central polypurine tract (cPPT)/central termination sequence (CTS) (e.g. DNA flap), a Poly A tail sequence, a posttranscriptional regulatory element (e.g. WPRE), a Rev response element (RRE), and a 3' LTR (e.g., comprising U5 and lacking a functional U3).

98. A method for selectively modulating the activity of CD8+ T cells, comprising contacting a viral vector according to any one of claims 56-96 with cells comprising CD8+ T cells.
99. A method of delivering an exogenous agent to a subject, comprising administering to the subject a viral vector according to any one of claims 56-96, wherein the viral vector further comprises an exogenous agent.
100. The method of claim 99, wherein the exogenous agent encodes a therapeutic agent or a diagnostic agent.
101. The method of claim 100, wherein the therapeutic agent is a chimeric antigen receptor (CAR) or T-cell receptor (TCR).
102. The method of claim 101, wherein the CAR targets a tumor antigen selected from CD19, CD20, CD22, or BCMA.
103. The method of claim 101 or 102, wherein the CAR comprises one or more intracellular domains selected from a CD137 (4-1BB) signaling domain, a CD28 signaling domain, and a CD3zeta signaling domain.
104. A method of treating cancer in a subject, comprising administering to the subject a viral vector according to any one of claims 56-97, wherein the viral vector further comprises an exogenous agent.
105. The method of claim 104, wherein the exogenous agent encodes a therapeutic agent or a diagnostic agent.
106. The method of claim 105, wherein the therapeutic agent is a chimeric antigen receptor (CAR) or T-cell receptor (TCR).
107. The method of claim 106, wherein the CAR targets a tumor antigen selected from CD19, CD20, CD22, or BCMA.
108. The method of claim 106 or 107, wherein the CAR comprises one or more intracellular domains selected from a CD137 (4-1BB) signaling domain, a CD28 signaling domain, and a CD3zeta signaling domain.
109. A method of transducing a cell that expresses CD8, comprising contacting the cell with the viral vector of any of claims 56-97.

110. The method of claim 109, wherein the cell is a CD8+ T cell.
111. The method of claim 109 or 110, wherein the viral vector comprises a chimeric antigen receptor (CAR) or T-cell receptor (TCR).
112. The method of claim 110 or 111, wherein the CD8+ T cell is not contacted with an exogenous activating agent.
113. Use of a viral vector according to any one of claims 56-97 for selectively modulating the activity of CD8+ T cells.
114. Use of a viral vector according to any one of claims 56-97 for selectively transducing CD8+ T cells.
115. A viral vector according to any one of claims 56-97 for use as a medicament.
116. Use of a viral vector according to any one of claims 56-97 for gene delivery.
117. The use of claim 89, wherein the viral vector is for delivery of a chimeric antigen receptor (CAR) or T-cell receptor (TCR).
118. Use of a viral vector according to any one of claims 56-97 for treating a B cell malignancy selected from the group consisting of Non-Hodgkin lymphoma (NHL), acute lymphocytic leukemia (ALL), chronic lymphocytic leukemia (CLL), and multiple myeloma.
119. The method of any of one of claims 98-112, wherein the contacting or administering is carried out by ex vivo administration of the lentiviral vector to a subject using a closed fluid circuit.
120. The method of claim 119, wherein the ex vivo administration comprises:
 - (a) obtaining whole blood from a subject;
 - (b) collecting the fraction of blood containing leukocyte components comprising T cells (e.g. CD8+ T cells);
 - (c) contacting the leukocyte components comprising T cells (e.g. CD8+ T cells) with a composition comprising the lentiviral vector; and
 - (d) reinfusing the contacted leukocyte components comprising T cells (e.g. CD8+ T cells) into the subject, wherein steps (a)-(d) are performed in-line in a closed fluid circuit.
121. The method of claim 120, wherein the contacting in step (c) is for not more than 24 hours, no more than 18 hours, no more than 12 hours, no more than 6 hours, no more than 4 hours, no more than 2 hours, or no more than 1 hour.