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

Innovation, Science and Economic Development Canada
Canadian Intellectual Property Office

CA 3149421 A1 2021/03/04

(21) 3 149 421

(12) DEMANDE DE BREVET CANADIEN
CANADIAN PATENT APPLICATION

(13) A1

(86) Date de dépôt PCT/PCT Filing Date: 2020/08/31
(87) Date publication PCT/PCT Publication Date: 2021/03/04
(85) Entrée phase nationale/National Entry: 2022/02/24
(86) N° demande PCT/PCT Application No.: US 2020/048823
(87) N° publication PCT/PCT Publication No.: 2021/042060
(30) Priorités/Priorities: 2019/08/30 (US PCT/US2019/048953); 2019/08/30 (US PCT/US2019/048962); 2019/12/05 (US62/944,281)

(51) Cl.Int./Int.Cl. *A61K 39/395* (2006.01), *A61K 48/00* (2006.01), *C07K 16/44* (2006.01)

(71) Demandeur/Applicant:
YALE UNIVERSITY, US

(72) Inventeurs/Inventors:
QUIJANO, ELIAS, US;
GLAZER, PETER, US

(74) Agent: BERESKIN & PARR LLP/S.E.N.C.R.L.,S.R.L.

(54) Titre : COMPOSITIONS ET METHODES D'ADMINISTRATION D'ACIDES NUCLEIQUES A DES CELLULES
(54) Title: COMPOSITIONS AND METHODS FOR DELIVERY OF NUCLEIC ACIDS TO CELLS

(57) Abrégé/Abstract:
Compositions and methods of use thereof for delivering nucleic acid cargo into cells are provided. The compositions typically include (a) a 3E10 monoclonal antibody or an antigen binding, cell-penetrating fragment thereof; a monovalent, divalent, or multivalent single chain variable fragment (scFv); or a diabody; or humanized form or variant thereof, and (b) a nucleic acid cargo including, for example, a nucleic acid encoding a polypeptide, a functional nucleic acid, a nucleic acid encoding a functional nucleic acid, or a combination thereof. Elements (a) and (b) are typically non-covalently linked to form a complex.

Date Submitted: 2022/02/24

CA App. No.: 3149421

Abstract:

Compositions and methods of use thereof for delivering nucleic acid cargo into cells are provided. The compositions typically include (a) a 3E10 monoclonal antibody or an antigen binding, cell-penetrating fragment thereof; a monovalent, divalent, or multivalent single chain variable fragment (scFv); or a diabody; or humanized form or variant thereof, and (b) a nucleic acid cargo including, for example, a nucleic acid encoding a polypeptide, a functional nucleic acid, a nucleic acid encoding a functional nucleic acid, or a combination thereof. Elements (a) and (b) are typically non-covalently linked to form a complex.

COMPOSITIONS AND METHODS FOR DELIVERY OF NUCLEIC ACIDS TO CELLS

CROSS-REFERENCE TO RELATED APPLICATIONS

5 This application claims the benefit of and priority to U.S. Provisional Application No. 62/944,281, entitled "*Compositions And Methods For Delivery Of Nucleic Acids To Cells*", filed in the United States Patent and Trademark Office on December 5, 2019, International Application No. PCT/US2019/048953, entitled "*Compositions And Methods For Enhancing* 10 *Donor Oligonucleotide-Based Gene Editing*" and filed in the United States Receiving Office for the Patent Cooperation Treaty on August 30, 2019, and International Application No. PCT/US2019/048962, entitled "*Compositions And Methods For Enhancing Triplex And Nuclease-Based Gene Editing*" and filed in the United States Receiving Office for the Patent Cooperation 15 Treaty on August 30, 2019. U.S. Provisional Application No. 62/944,281, International Application No. PCT/US2019/048953, U.S. Provisional Application No. 62/725,920, International Application No. PCT/US2019/048962, U.S. Provisional Application No. 62/725,852 are each specifically incorporated by reference in their entireties.

20

STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH

This invention was made with government support under CA197574 awarded by National Institutes of Health. The Government has certain rights in the invention.

25

REFERENCE TO THE SEQUENCE LISTING

The Sequence Listing submitted as a text file named "YU_7503_3_ST25" created on August 31, 2020, and having a size of 154,701 bytes is hereby incorporated by reference pursuant to 37 C.F.R. § 1.52(e)(5).

30

FIELD OF THE INVENTION

The invention is generally related to the field of intracellular delivery of nucleic acids, for application including, but not limited to *in vitro*, *ex vivo*, and *in vivo* gene therapy and gene editing.

CA 03149421 2022-2-24

BACKGROUND OF THE INVENTION

Gene therapy includes a spectrum of applications ranging from gene replacement and knockdown for genetic or acquired diseases such as cancer, to vaccination. Viral vectors and synthetic liposomes have emerged as the
5 vehicles of choice for many applications today, but both have limitations and risks, including complexity of production, limited packaging capacity, and unfavorable immunological features, which restrict gene therapy applications and hold back the potential for preventive gene therapy (Seow and Wood, *Mol Ther.* 17(5): 767–777 (2009).

10 *In vivo* uptake and distributed of nucleotide in cells and tissues has been observed (Huang, et al., *FEBS Lett.,* 558(1-3):69-73 (2004)). Further, although, for example, Nyce, et al. have shown that antisense oligodeoxynucleotides (ODNs) when inhaled bind to endogenous surfactant (a lipid produced by lung cells) and are taken up by lung cells without a need
15 for additional carrier lipids (Nyce, et al., *Nature*, 385:721-725 (1997)), small nucleic acids are taken up into T24 bladder carcinoma tissue culture cells (Ma, et al., *Antisense Nucleic Acid Drug Dev.,* 8:415-426 (1998)), there remains a need for improved nucleic acid transfection technology, particularly for *in vivo* applications. AAV9, still the viral vector typically
20 used in people was discovered in 2003 (Robbins, "Gene therapy pioneer says the field is behind – and that delivery technology is embarrassing," *Stat*, November, 2019).

Thus, it is an object of the invention to provided compositions and methods of use thereof for improved delivery of nucleic acids into cells.

25

SUMMARY OF THE INVENTION

Compositions and methods of use thereof for delivering nucleic acid cargo into cells are provided. The compositions typically include (a) a 3E10 monoclonal antibody or a cell-penetrating fragment thereof; a monovalent, divalent, or multivalent single chain variable fragment (scFv); or a diabody;
30 or humanized form or variant thereof, and (b) a nucleic acid cargo including, for example, a nucleic acid encoding a polypeptide, a functional nucleic acid, a nucleic acid encoding a functional nucleic acid, or a combination thereof. Elements (a) and (b) are typically non-covalently linked to form a complex.

CA 03149421 2022-2-24

It is believed that in additional to DNA, 3E10 binds to RNA, PNA, and other nucleic acids.

Exemplary 3E10 antibodies and fragments and fusion protein thereof include those having (i) the CDRs of any one of SEQ ID NO:1-6, 12, 13, 46-48, or 50-52 in combination with the CDRs of any one of SEQ ID NO:7-11, 14, or 53-58; (ii) first, second, and third heavy chain CDRs selected from SEQ ID NOS:15-23, 42, and 43 in combination with first, second and third light chain CDRs selected from SEQ ID NOS:24-30, 44, and 45; (iii) a humanized forms of (i) or (ii); (iv) a heavy chain comprising an amino acid sequence comprising at least 85% sequence identity to any one of SEQ ID NO:1 or 2 in combination with a light chain comprising an amino acid sequence comprising at least 85% sequence identity to SEQ ID NO:7 or 8; (v) a humanized form or (iv); or (vi) a heavy chain comprising an amino acid sequence comprising at least 85% sequence identity to any one of SEQ ID NO:3-6, 46-48, or 50-52 in combination with a light chain comprising an amino acid sequence comprising at least 85% sequence identity to SEQ ID NO:9-11 or 53-58.

In some embodiments, the antibodies and fragments and fusion protein thereof is CDR1 heavy chain variant having the amino acid residue corresponding with D31 or N31 of a 3E10 heavy chain amino acid sequence or a CDR thereof substituted with arginine (R) or lysine (L).

In some embodiments, the antibodies and fragments and fusion protein thereof include the nucleic acid binding pocket of SEQ ID NOS:92 or 93, or a variant thereof with same or improved ability to bind to a nucleic acid, such as DNA, RNA, or a combination thereof.

Also provided are binding proteins themselves including a CDR1 heavy chain variant having the amino acid residue corresponding with D31 or N31 of a 3E10 heavy chain amino acid sequence or the CDR1 thereof substituted with arginine (R) or lysine (L), as well as binding proteins themselves having the nucleic acid binding pocket of SEQ ID NOS:92 or 93, or a variant thereof with same or improved ability to bind to a nucleic acid, such as DNA, RNA, or a combination thereof.

CA 03149421 2022-2-24

In some embodiments, the antibody or fragment or fusion protein can be bispecific, and can, for example, include a binding sequence that targets a cell type, tissue, or organ of interest.

The nucleic acid cargo can be composed of DNA, RNA, modified
5 nucleic acids, including but not limited to, PNA, or a combination thereof. The nucleic acid cargo is typically a functional cargo, such as a functional nucleic (e.g., an inhibitory RNA), an mRNA, or a vector, for example an expression vector. The nucleic acid cargo, including vectors, can include a nucleic acid sequence encoding a polypeptide of interest operably linked to
10 expression control sequence. The vector can be, for example, a plasmid. Typically the cargo is not, for example, randomly sheared or fragment genomic DNA.

In some embodiments, the cargo includes or consists of a nucleic acid encoding a Cas endonuclease, a gRNA, or a combination thereof. In some
15 embodiments, the cargo includes or consists of a nucleic acid encoding a chimeric antigen receptor polypeptide. In some embodiments, the cargo is a functional nucleic acid such as antisense molecules, siRNA, microRNA (miRNA), aptamers, ribozymes, RNAi, or external guide sequences, or a nucleic acid construct encoding the same.

20 The cargo can include or consist of a plurality of a single nucleic acid molecule, or a plurality of 2, 3, 4, 5, 6, 7, 8, 9, 10, or more different nucleic acid molecules. In some embodiments, the nucleic acid molecules of cargo include or consists of nucleic acid molecules between about 1 and about 25,000 nucleobases in length. The cargo can be single stranded nucleic
25 acids, double stranded nucleic acids, or a combination thereof.

Pharmaceutical compositions including the complexes and a pharmaceutically acceptable excipient are also provided. In some embodiments, the complexes are encapsulated in polymeric nanoparticles. A targeting moiety, a cell-penetrating peptide, or a combination thereof can be
30 associated with, linked, conjugated, or otherwise attached directly or indirectly to the nanoparticle.

Methods of delivering into cells, the nucleic acid cargo, by contacting the cells with an effective amount of the complexes alone or encapsulated in nanoparticles are also provided. The contacting can occur in vitro, ex vivo,

CA 03149421 2022-2-24

or in vivo. In some embodiments, an effective amount of ex vivo treated cells are administered to a subject in need thereof, e.g., in an effective amount to treat one or more symptoms of a disease or disorder.

In some embodiments, the contacting occurs in vivo following administration to a subject in need thereof. The subject can have a disease or disorder, such as a genetic disorder or cancer. The compositions can be administered to the subject, for example by injection or infusion, in an effective amount to reduce one or more symptoms of the disease or disorder in the subject.

Applications of the compositions and methods are also provided, and include, but are not limited to, gene therapy and CAR T cell manufacture/formation/therapy.

BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A-1C are scatter plots showing control (1A), and uptake of PNA when alone (1B) and when mixed with 3E10 for 1 hour (1C). Figure 1D is a bar graph quantifying the data in Figures 1A-1C.

Figures 2A-2C are scatter plots showing control (2A), and uptake of PNA when alone (2B) and when mixed with 3E10 for 24 hour (2C). Figure 2D is a bar graph quantifying the data in Figures 2A-2C.

Figures 3A-3C are scatter plots showing control (3A), and uptake of siRNA when alone (3B) and when mixed with 3E10 for 24 hour (3C). Figure 3D is a bar graph quantifying the data in Figures 3A-3C.

Figures 4A-4H are scatter plots showing control (4A), and uptake of mRNA when alone (4B) and when mixed at various concentrations with 3E10 for 24 hour (4C-4H). Figure 4I is a bar graph quantifying the data in Figures 4A-4H.

Figures 5A-5H are scatter plots showing control (5A), and uptake of mRNA when alone (5B) and when mixed at various concentrations with 3E10 for 1 hour (5C-5H). Figure 5I is a bar graph quantifying the data in Figures 5A-5H.

Figure 6 is a series of images showing cellular expression of a GFP reporter plasmid DNA 72 hours after mixture with 3E10 and 24 hours of incubation with cells.

CA 03149421 2022-2-24

Figure 7A is a bar graph showing accumulation in tumors of fluorescently labeled siRNA mixed with increasing doses of 3E10 (0.25, 0.5, and 1 mg) for 15 minutes at room temperature prior to systemic injection in mice. Figure 7B is a bar graph showing accumulation in tumors of fluorescently labeled siRNA mixed with 1 mg 3E10 or 0.1 mg D31N variant 3E10 for 15 minutes at room temperature prior to systemic injection in mice. All tumors were analyzed 24 hours after injection.

Figure 8 is a line graph showing 3E10-mediated delivery of mRNA (bioluminescene (Photons/second)) to mouse muscles (IM) over time (days post-IM injection).

Figures 9A and 9B are images showing control (Fig. 9A) and distribution of IV Injected 3E10-D31N to muscle (Fig. 9B), imaged by IVIS (Perkin Elmer) 24 hours after injection. Figure 9C is a bar graph quantifying the fluorescence in the IVIS images.

Figure 10 is a bar graph quantifying the fluorescence in the IVIS images of dose-dependent biodistribution of 3E10-D31N to tissues 24 hours following 100 μg or 200 μg intravenous injection of 3E10-D31N labeled with VivoTag680 into mice (Perkin Elmer).

Figures 11A and 11B are images showing control (Fig. 11A) and distribution of IV Injected 3E10-D31N to syngeneic colon tumors (CT26) (Fig. 11B), imaged by IVIS (Perkin Elmer) 24 hours after injection. Figure 11C is a bar graph quantifying the fluorescence in the IVIS images.

Figures 12A, 12B, and 12C are images showing control (Fig. 12A), and distribution of IV Injected naked single stranded DNA (ssDNA) (Fig. 12B) and 3E10-D31N + ssDNA (Fig. 12C) syngeneic colon tumors (CT26), imaged by IVIS (Perkin Elmer) 24 hours after injection. Figure 12D is a bar graph quantifying the fluorescence in the IVIS images.

Figure 13 is a bar graph showing 3E10-mediated delivery and stimulation of RIG-I.

Figure 14A is an illustration of molecular modeling of 3E10, a putative Nucleic Acid Binding pocket (NAB1) thereof, and the predicted structural changes induced by amino acid mutations therein. Figure 14B is an illustration of molecular modeling of 3E10-scFv (Pymol) with NAB1 amino acid residues highlighted by punctate dots.

CA 03149421 2022-2-24

DETAILED DESCRIPTION OF THE INVENTION

I. Definitions

As used herein, the term "single chain Fv" or "scFv" as used herein means a single chain variable fragment that includes a light chain variable region (VL) and a heavy chain variable region (VH) in a single polypeptide chain joined by a linker which enables the scFv to form the desired structure for antigen binding (i.e., for the VH and VL of the single polypeptide chain to associate with one another to form a Fv). The VL and VH regions may be derived from the parent antibody or may be chemically or recombinantly synthesized.

As used herein, the term "variable region" is intended to distinguish such domain of the immunoglobulin from domains that are broadly shared by antibodies (such as an antibody Fc domain). The variable region includes a "hypervariable region" whose residues are responsible for antigen binding. The hypervariable region includes amino acid residues from a "Complementarity Determining Region" or "CDR" (i.e., typically at approximately residues 24-34 (L1), 50-56 (L2) and 89-97 (L3) in the light chain variable domain and at approximately residues 27-35 (H1), 50-65 (H2) and 95-102 (H3) in the heavy chain variable domain; Kabat et al., Sequences of Proteins of Immunological Interest, 5th Ed. Public Health Service, National Institutes of Health, Bethesda, MD. (1991)) and/or those residues from a "hypervariable loop" (i.e., residues 26-32 (L1), 50-52 (L2) and 91-96 (L3) in the light chain variable domain and 26-32 (H1), 53-55 (H2) and 96-101 (H3) in the heavy chain variable domain; Chothia and Lesk, 1987, *J. Mol. Biol.* 196:901-917).

As used herein, the term "Framework Region" or "FR" residues are those variable domain residues other than the hypervariable region residues as herein defined.

As used herein, the term "antibody" refers to natural or synthetic antibodies that bind a target antigen. The term includes polyclonal and monoclonal antibodies. In addition to intact immunoglobulin molecules, also included in the term "antibodies" are binding proteins, fragments, and polymers of those immunoglobulin molecules, and human or humanized versions of immunoglobulin molecules that bind the target antigen.

CA 03149421 2022-2-24

As used herein, the term "cell-penetrating antibody" refers to an immunoglobulin protein, fragment, variant thereof, or fusion protein based thereon that is transported into the cytoplasm and/or nucleus of living mammalian cells. The "cell-penetrating anti-DNA antibody" specifically binds DNA (*e.g.*, single-stranded and/or double-stranded DNA). In some embodiments, the antibody is transported into the cytoplasm of the cells without the aid of a carrier or conjugate. In other embodiments, the antibody is conjugated to a cell-penetrating moiety, such as a cell penetrating peptide. In some embodiments, the cell-penetrating antibody is transported in the nucleus with or without a carrier or conjugate.

In addition to intact immunoglobulin molecules, also included in the term "antibodies" are fragments, binding proteins, and polymers of immunoglobulin molecules, chimeric antibodies containing sequences from more than one species, class, or subclass of immunoglobulin, such as human or humanized antibodies, and recombinant proteins containing a least the idiotype of an immunoglobulin that specifically binds DNA. The antibodies can be tested for their desired activity using the *in vitro* assays described herein, or by analogous methods, after which their *in vivo* therapeutic activities are tested according to known clinical testing methods.

As used herein, the term "variant" refers to a polypeptide or polynucleotide that differs from a reference polypeptide or polynucleotide, but retains essential properties. A typical variant of a polypeptide differs in amino acid sequence from another, reference polypeptide. Generally, differences are limited so that the sequences of the reference polypeptide and the variant are closely similar overall and, in many regions, identical. A variant and reference polypeptide may differ in amino acid sequence by one or more modifications (*e.g.*, substitutions, additions, and/or deletions). A substituted or inserted amino acid residue may or may not be one encoded by the genetic code. A variant of a polypeptide may be naturally occurring such as an allelic variant, or it may be a variant that is not known to occur naturally.

Modifications and changes can be made in the structure of the polypeptides of in disclosure and still obtain a molecule having similar characteristics as the polypeptide (*e.g.*, a conservative amino acid

CA 03149421 2022-2-24

substitution). For example, certain amino acids can be substituted for other amino acids in a sequence without appreciable loss of activity. Because it is the interactive capacity and nature of a polypeptide that defines that polypeptide's biological functional activity, certain amino acid sequence substitutions can be made in a polypeptide sequence and nevertheless obtain a polypeptide with like properties.

In making such changes, the hydropathic index of amino acids can be considered. The importance of the hydropathic amino acid index in conferring interactive biologic function on a polypeptide is generally understood in the art. It is known that certain amino acids can be substituted for other amino acids having a similar hydropathic index or score and still result in a polypeptide with similar biological activity. Each amino acid has been assigned a hydropathic index on the basis of its hydrophobicity and charge characteristics. Those indices are: isoleucine (+4.5); valine (+4.2); leucine (+3.8); phenylalanine (+2.8); cysteine/cystine (+2.5); methionine (+1.9); alanine (+1.8); glycine (-0.4); threonine (-0.7); serine (-0.8); tryptophan (-0.9); tyrosine (-1.3); proline (-1.6); histidine (-3.2); glutamate (-3.5); glutamine (-3.5); aspartate (-3.5); asparagine (-3.5); lysine (-3.9); and arginine (-4.5).

It is believed that the relative hydropathic character of the amino acid determines the secondary structure of the resultant polypeptide, which in turn defines the interaction of the polypeptide with other molecules, such as enzymes, substrates, receptors, antibodies, antigens, and cofactors. It is known in the art that an amino acid can be substituted by another amino acid having a similar hydropathic index and still obtain a functionally equivalent polypeptide. In such changes, the substitution of amino acids whose hydropathic indices are within ± 2 is preferred, those within ± 1 are particularly preferred, and those within ± 0.5 are even more particularly preferred.

Substitution of like amino acids can also be made on the basis of hydrophilicity, particularly where the biological functional equivalent polypeptide or peptide thereby created is intended for use in immunological embodiments. The following hydrophilicity values have been assigned to amino acid residues: arginine (+3.0); lysine (+3.0); aspartate (+3.0 ± 1);

glutamate (+3.0 ± 1); serine (+0.3); asparagine (+0.2); glutamnine (+0.2); glycine (0); proline (-0.5 ± 1); threonine (-0.4); alanine (-0.5); histidine (-0.5); cysteine (-1.0); methionine (-1.3); valine (-1.5); leucine (-1.8); isoleucine (-1.8); tyrosine (-2.3); phenylalanine (-2.5); tryptophan (-3.4). It
5 is understood that an amino acid can be substituted for another having a similar hydrophilicity value and still obtain a biologically equivalent, and in particular, an immunologically equivalent polypeptide. In such changes, the substitution of amino acids whose hydrophilicity values are within ± 2 is preferred, those within ± 1 are particularly preferred, and those within ± 0.5
10 are even more particularly preferred.

As outlined above, amino acid substitutions are generally based on the relative similarity of the amino acid side-chain substituents, for example, their hydrophobicity, hydrophilicity, charge, size, and the like. Exemplary substitutions that take various of the foregoing characteristics into
15 consideration are well known to those of skill in the art and include (original residue: exemplary substitution): (Ala: Gly, Ser), (Arg: Lys), (Asn: Gln, His), (Asp: Glu, Cys, Ser), (Gln: Asn), (Glu: Asp), (Gly: Ala), (His: Asn, Gln), (Ile: Leu, Val), (Leu: Ile, Val), (Lys: Arg), (Met: Leu, Tyr), (Ser: Thr), (Thr: Ser), (Tip: Tyr), (Tyr: Trp, Phe), and (Val: Ile, Leu). Embodiments of
20 this disclosure thus contemplate functional or biological equivalents of a polypeptide as set forth above. In particular, embodiments of the polypeptides can include variants having about 50%, 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to the polypeptide of interest.

25 As used herein, the term "percent (%) sequence identity" is defined as the percentage of nucleotides or amino acids in a candidate sequence that are identical with the nucleotides or amino acids in a reference nucleic acid sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity. Alignment for purposes of
30 determining percent 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, ALIGN-2 or Megalign (DNASTAR) software. Appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full-

CA 03149421 2022-2-24

length of the sequences being compared can be determined by known methods.

For purposes herein, the % sequence identity of a given nucleotides or amino acids sequence C to, with, or against a given nucleic acid sequence D (which can alternatively be phrased as a given sequence C that has or includes a certain % sequence identity to, with, or against a given sequence D) is calculated as follows:

100 times the fraction W/Z,

where W is the number of nucleotides or amino acids scored as identical matches by the sequence alignment program in that program's alignment of C and D, and where Z is the total number of nucleotides or amino acids in D. It will be appreciated that where the length of sequence C is not equal to the length of sequence D, the % sequence identity of C to D will not equal the % sequence identity of D to C.

As used herein, the term "specifically binds" refers to the binding of an antibody to its cognate antigen (for example, DNA) while not significantly binding to other antigens. Specific binding of an antibody to a target under such conditions requires the antibody be selected for its specificity to the target. A variety of immunoassay formats may be used to select antibodies specifically immunoreactive with a particular protein. For example, solid-phase ELISA immunoassays are routinely used to select monoclonal antibodies specifically immunoreactive with a protein. See, e.g., Harlow and Lane (1988) Antibodies, A Laboratory Manual, Cold Spring Harbor Publications, New York, for a description of immunoassay formats and conditions that can be used to determine specific immunoreactivity. Preferably, an antibody "specifically binds" to an antigen with an affinity constant (Ka) greater than about 10^5 mol^{-1} (e.g., 10^6 mol^{-1}, 10^7 mol^{-1}, 10^8 mol^{-1}, 10^9 mol^{-1}, 10^{10} mol^{-1}, 10^{11} mol^{-1}, and 10^{12} mol^{-1} or more) with that second molecule.

As used herein, the term "monoclonal antibody" or "MAb" refers to an antibody obtained from a substantially homogeneous population of antibodies, i.e., the individual antibodies within the population are identical except for possible naturally occurring mutations that may be present in a small subset of the antibody molecules.

CA 03149421 2022-2-24

As used herein, the term "subject" means any individual who is the target of administration. The subject can be a vertebrate, for example, a mammal. Thus, the subject can be a human. The term does not denote a particular age or sex.

5 As used herein, the term "effective amount" means that the amount of the composition used is of sufficient quantity to ameliorate one or more causes or symptoms of a disease or disorder. Such amelioration only requires a reduction or alteration, not necessarily elimination. The precise dosage will vary according to a variety of factors such as subject-dependent 10 variables (e.g., age, immune system health, etc.), the disease or disorder being treated, as well as the route of administration and the pharmacokinetics of the agent being administered.

As used herein, the term "pharmaceutically acceptable" refers to a material that is not biologically or otherwise undesirable, i.e., the material 15 may be administered to a subject without causing any undesirable biological effects or interacting in a deleterious manner with any of the other components of the pharmaceutical composition in which it is contained.

As used herein, the term "carrier" or "excipient" refers to an organic or inorganic ingredient, natural or synthetic inactive ingredient in a 20 formulation, with which one or more active ingredients are combined. The carrier or excipient would naturally be selected to minimize any degradation of the active ingredient and to minimize any adverse side effects in the subject, as would be well known to one of skill in the art.

As used herein, the term "treat" refers to the medical management of 25 a patient with the intent to cure, ameliorate, stabilize, or prevent a disease, pathological condition, or disorder. This term includes active treatment, that is, treatment directed specifically toward the improvement of a disease, pathological condition, or disorder, and also includes causal treatment, that is, treatment directed toward removal of the cause of the associated disease, 30 pathological condition, or disorder. In addition, this term includes palliative treatment, that is, treatment designed for the relief of symptoms rather than the curing of the disease, pathological condition, or disorder; preventative treatment, that is, treatment directed to minimizing or partially or completely inhibiting the development of the associated disease, pathological condition,

CA 03149421 2022-2-24

or disorder; and supportive treatment, that is, treatment employed to supplement another specific therapy directed toward the improvement of the associated disease, pathological condition, or disorder.

As used herein, "targeting moiety" is a substance which can direct a particle or molecule to a receptor site on a selected cell or tissue type, can serve as an attachment molecule, or serve to couple or attach another molecule. As used herein, "direct" refers to causing a molecule to preferentially attach to a selected cell or tissue type. This can be used to direct cellular materials, molecules, or drugs, as discussed below.

As used herein, the term "inhibit" or "reduce" means to decrease an activity, response, condition, disease, or other biological parameter. This can include, but is not limited to, the complete ablation of the activity, response, condition, or disease. This may also include, for example, a 10% reduction in the activity, response, condition, or disease as compared to the native or control level. Thus, the reduction can be a 10, 20, 30, 40, 50, 60, 70, 80, 90, 100%, or any amount of reduction in between as compared to native or control levels.

As used herein, a "fusion protein" refers to a polypeptide formed by the joining of two or more polypeptides through a peptide bond formed between the amino terminus of one polypeptide and the carboxyl terminus of another polypeptide. The fusion protein can be formed by the chemical coupling of the constituent polypeptides or it can be expressed as a single polypeptide from a nucleic acid sequence encoding the single contiguous fusion protein. A single chain fusion protein is a fusion protein having a single contiguous polypeptide backbone. Fusion proteins can be prepared using conventional techniques in molecular biology to join the two genes in frame into a single nucleic acid sequence, and then expressing the nucleic acid in an appropriate host cell under conditions in which the fusion protein is produced.

Recitation of ranges of values herein are merely intended to serve as a shorthand method of referring individually to each separate value falling within the range, unless otherwise indicated herein, and each separate value is incorporated into the specification as if it were individually recited herein.

Use of the term “about” is intended to describe values either above or below the stated value in a range of approx. +/- 10%; in other embodiments the values may range in value either above or below the stated value in a range of approx. +/- 5%; in other embodiments the values may range in value either above or below the stated value in a range of approx. +/- 2%; in other embodiments the values may range in value either above or below the stated value in a range of approx. +/- 1%. The preceding ranges are intended to be made clear by context, and no further limitation is implied.

All methods described herein can be performed in any suitable order unless otherwise indicated or otherwise clearly contradicted by context. The use of any and all examples, or exemplary language (e.g., “such as”) provided herein, is intended merely to better illuminate the embodiments and does not pose a limitation on the scope of the embodiments unless otherwise claimed. No language in the specification should be construed as indicating any non-claimed element as essential to the practice of the invention.

II. Compositions

It has been discovered that 3E10 antibody helps deliver nucleic acids across the plasma membrane and into cell cytoplasm and nuclei. Thus, compositions and methods for using 3E10 to enhance delivery of nucleic acid constructs are provided. Typically an effective amount of 3E10 antibody is contacted with a nucleic acid whose delivery into cells is desired. Typically, the contacting occurs for a sufficient amount to time for the 3E10 and the nucleic acid cargo to form a complex. The complexes are contacted with cells for a sufficient amount of time for the nucleic acid cargo to be delivered into the cells. The cargo may accumulate in a greater quantity, greater quality (e.g., more intact, functional, etc.), or a faster rate, or combination thereof, than if the cells were contacted with the nucleic acid cargo in the absence of the antibody. Because the antibody serves as the delivery means, the delivery systems are typically non-viral.

A. 3E10 Antibodies

Although generally referred to herein as “3E10” or “3E10 antibodies,” it will be appreciated that fragments and binding proteins, including antigen-binding fragments, variants, and fusion proteins such as scFv, di-scFv, tr-scFv, and other single chain variable fragments, and other

CA 03149421 2022-2-24

cell-penetrating, nucleic acid transporting molecules disclosed herein are encompassed by the phrase are also expressly provided for use in compositions and methods disclosed herein. Thus, the antibodies and other binding proteins are also referred to herein as cell-penetrating.

5 In preferred embodiments, the 3E10 antibody is transported into the cytoplasm and/or nucleus of the cells without the aid of a carrier or conjugate. For example, the monoclonal antibody 3E10 and active fragments thereof that are transported *in vivo* to the nucleus of mammalian cells without cytotoxic effect are disclosed in U.S. Patent Nos. 4,812,397 and 7,189,396 to 10 Richard Weisbart.

In some embodiments, the antibody may bind and/or inhibit Rad51. See for example, the antibody described in Turchick, et al., *Nucleic Acids Res.*, 45(20): 11782-11799 (2017), WO 2020/047344, and WO 2020/047353, each of which is specifically incorporated by reference herein, in its entirety.

15 Antibodies that can be used in the compositions and methods include whole immunoglobulin (i.e., an intact antibody) of any class, fragments thereof, and synthetic proteins containing at least the antigen binding variable domain of an antibody. The variable domains differ in sequence among antibodies and are used in the binding and specificity of each 20 particular antibody for its particular antigen. However, the variability is not usually evenly distributed through the variable domains of antibodies. It is typically concentrated in three segments called complementarity determining regions (CDRs) or hypervariable regions both in the light chain and the heavy chain variable domains. The more highly conserved portions of the 25 variable domains are called the framework (FR). The variable domains of native heavy and light chains each comprise four FR regions, largely adopting a beta-sheet configuration, connected by three CDRs, which form loops connecting, and in some cases forming part of, the beta-sheet structure. The CDRs in each chain are held together in close proximity by the FR 30 regions and, with the CDRs from the other chain, contribute to the formation of the antigen binding site of antibodies. Therefore, the antibodies typically contain at least the CDRs necessary to maintain DNA binding and/or interfere with DNA repair.

CA 03149421 2022-2-24

The 3E10 antibody is typically a monoclonal 3E10, or a variant, derivative, fragment, fusion, or humanized form thereof that binds the same or different epitope(s) as 3E10.

A deposit according to the terms of the Budapest Treaty of a hybridoma cell line producing monoclonal antibody 3E10 was received on September 6, 2000, and accepted by, American Type Culture Collection (ATCC), 10801 University Blvd., Manassas, VA 20110-2209, USA, and given Patent Deposit Number PTA-2439.

Thus, the antibody may have the same or different epitope specificity as monoclonal antibody 3E10 produced by ATCC No. PTA 2439 hybridoma. The antibody can have the paratope of monoclonal antibody 3E10. The antibody can be a single chain variable fragment of 3E10, or a variant, e.g., a conservative variant thereof. For example, the antibody can be a single chain variable fragment of 3E10 (3E10 Fv), or a variant thereof.

1. 3E10 Sequences

Amino acid sequences of monoclonal antibody 3E10 are known in the art. For example, sequences of the 3E10 heavy and light chains are provided below, where single underlining indicates the CDR regions identified according to the Kabat system, and in SEQ ID NOS:12-14 italics indicates the variable regions and double underlining indicates the signal peptide. CDRs according to the IMGT system are also provided.

a. 3E10 Heavy Chain

In some embodiments, a heavy chain variable region of 3E10 is:
EVQLVESGGGLVKPGGSRKLSCAASGFTFS**D**YGMHWVRQAPEKGLEWVAYI
SSGSSTIYYADTVKGRFTISRDNAKNTLFLQMTSLRSEDTAMYYCARRGLL
LDYWGQGTTLTVSS (SEQ ID NO:1; Zack, et al., *Immunology and Cell Biology*, 72:513-520 (1994); GenBank: L16981.1 - *Mouse Ig rearranged L-chain gene, partial cds*; and GenBank: AAA65679.1 - *immunoglobulin heavy chain, partial [Mus musculus]*).

In some embodiments, a 3E10 heavy chain is expressed as
MGWSCIILFLVATATGVHS*EVQLVESGGGLVKPGGSRKLSCAASGFTFS***D***Y*
GMHWVRQAPEKGLEWVAYISSGSSTIYYADTVKGRFTISRDNAKNTLFLQM
*TSLRSEDTAMYYCARRGLLLDYWGQGTTLTVS*AASTKGPSVFPLAPSSKST
SGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVV

CA 03149421 2022-2-24

TVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTCPPCPAPELLGG
PSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAK
TKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA
KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENN
YKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSL
SLSPGK (3E10 WT Heavy Chain; SEQ ID NO:12).

Variants of the 3E10 antibody which incorporate mutations into the wild type sequence are also known in the art, as disclosed for example, in Zack, et al., *J. Immunol.*, 157(5):2082-8 (1996). For example, amino acid position 31 of the heavy chain variable region of 3E10 has been determined to be influential in the ability of the antibody and fragments thereof to penetrate nuclei and bind to DNA (bolded in SEQ ID NOS:1, 2 and 13). A D31N mutation (bolded in SEQ ID NOS:2 and 13) in CDR1 penetrates nuclei and binds DNA with much greater efficiency than the original antibody (Zack, et al., *Immunology and Cell Biology*, 72:513-520 (1994), Weisbart, et al., *J. Autoimmun.*, 11, 539-546 (1998); Weisbart, *Int. J. Oncol.*, 25, 1867-1873 (2004)). In some embodiments, the antibody has the D31N substitution.

In some embodiments, an amino acid sequence for a preferred variant of a heavy chain variable region of 3E10 is:
EVQLVESGGGLVKPGGSRKLSCAASGFTFS**N**YGMHWVRQAPEKGLEWVAYI
SSGSSTIYYADTVKGRFTISRDNAKNTLFLQMTSLRSEDTAMYYCARRGLL
LDYWGQGTTLTVSS (SEQ ID NO:2).

In some embodiments, a 3E10 heavy chain is expressed as
MGWSCIILFLVATATGVHS*EVQLVESGGGLVKPGGSRKLSCAASGFTFS***N***Y*
GMHWVRQAPEKGLEWVAYISSGSSTIYYADTVKGRFTISRDNAKNTLFLQM
*TSLRSEDTAMYYCARRGLLLDYWGQGTTLTVS*AASTKGPSVFPLAPSSKST
SGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVV
TVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTCPPCPAPELLGG
PSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAK
TKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA
KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENN
YKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSL
SLSPGK (3E10 D31N Variant Heavy Chain; SEQ ID NO:13).

CA 03149421 2022-2-24

In some embodiments, the C-terminal serine of SEQ ID NOS:1 or 2 is absent or substituted, with, for example, an alanine, in 3E10 heavy chain variable region.

The complementarity determining regions (CDRs) as identified by Kabat are shown with underlining above and include CDR H1.1 (original sequence): DYGMH (SEQ ID NO:15); CDR H1.2 (with D31N mutation): NYGMH (SEQ ID NO:16); CDR H2.1: YISSGSSTIYYADTVKG (SEQ ID NO:17); CDR H3.1: RGLLLDY (SEQ ID NO:18).

Variants of Kabat CDR H2.1 include YISSGSSTIYYADSVKG (SEQ ID NO:19) and YISSSSSTIYYADSVKG (SEQ ID NO:42).

Additionally, or alternatively, the heavy chain complementarity determining regions (CDRs) can be defined according to the IMGT system. The complementarity determining regions (CDRs) as identified by the IMGT system include CDR H1.3 (original sequence): GFTFSDYG (SEQ ID NO:20); CDR H1.4 (with D31N mutation): GFTFSNYG (SEQ ID NO:21); CDR H2.2: ISSGSSTI (SEQ ID NO:22) and variant ISSSSSTI (SEQ ID NO:43); CDR H3.2: ARRGLLLDY (SEQ ID NO:23).

b. 3E10 Light Chain

In some embodiments, a light chain variable region of 3E10 is:

DIVLTQSPASLAVSLGQRATISCRASKSVSTSSYSYMHWYQQKPGQPPKLLIKYASYLESGVPARFSGSGSGTDFTLNIHPVEEEDAATYYCQHSREFPWTFGGGTKLEIK (SEQ ID NO:7).

An amino acid sequence for the light chain variable region of 3E10 can also be:

DIVLTQSPASLAVSLGQRATISCRASKSVSTSSYSYMHWYQQKPGQPPKLLIKYASYLESGVPARFSGSGSGTDFHLNIHPVEEEDAATYYCQHSREFPWTFGGGTKLELK (SEQ ID NO:8).

In some embodiments, a 3E10 light chain is expressed as MGWSCIILFLVATATGVHS*DIVLTQSPASLAVSLGQRATISCRASKSVSTSSYSYMHWYQQKPGQPPKLLIKYASYLESGVPARFSGSGSGTDFTLNIHPVEEEDAATYYCQHSREFPWTFGGGTKLEIKR*TVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTL

CA 03149421 2022-2-24

SKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC (3E10 WT Light Chain; SEQ ID NO:14)

Other 3E10 light chain sequences are known in the art. See, for example, Zack, et al., *J. Immunol.*, 15;154(4):1987-94 (1995); GenBank: L16981.1 - *Mouse Ig rearranged L-chain gene, partial cds*; GenBank: AAA65681.1 - *immunoglobulin light chain, partial [Mus musculus]*).

The complementarity determining regions (CDRs) as identified by Kabat are shown with underlining, including CDR L1.1: RASKSVSTSSYSYMH (SEQ ID NO:24); CDR L2.1: YASYLES (SEQ ID NO:25); CDR L3.1: QHSREFPWT (SEQ ID NO:26).

Variants of Kabat CDR L1.1 include RASKSVSTSSYSYLA (SEQ ID NO:27) and RASKTVSTSSYSYMH (SEQ ID NO:44).

A variant of Kabat CDR L2.1 is YASYLQS (SEQ ID NO:28).

Additionally, or alternatively, the heavy chain complementarity determining regions (CDRs) can be defined according to the IMGT system. The complementarity determining regions (CDRs) as identified by the IMGT system include CDR L1.2 KSVSTSSYSY (SEQ ID NO:29) and variant KTVSTSSYSY (SEQ ID NO:45); CDR L2.2: YAS (SEQ ID NO:30); CDR L3.2: QHSREFPWT (SEQ ID NO:26).

In some embodiments, the C-terminal end of sequence of SEQ ID NOS:7 or 8 further includes an arginine in the 3E10 light chain variable region.

2. Humanized 3E10

In some embodiments, the antibody is a humanized antibody. Methods for humanizing non-human antibodies are well known in the art. Generally, a humanized antibody has one or more amino acid residues introduced into it from a source that is non-human. These non-human amino acid residues are often referred to as "import" residues, which are typically taken from an "import" variable domain. Antibody humanization techniques generally involve the use of recombinant DNA technology to manipulate the DNA sequence encoding one or more polypeptide chains of an antibody molecule.

CA 03149421 2022-2-24

Exemplary 3E10 humanized sequences are discussed in WO 2015/106290, WO 2016/033324, WO 2019/018426, and WO/2019/018428, and provided below.

5

a. Humanized 3E10 Heavy Chain Variable Regions

In some embodiments, a humanized 3E10 heavy chain variable domain includes
EVQLVQSGGGLIQPGGSLRLSCAASGFTFSNYGMHWVRQAPGKGLEWVSYI
SSGSSTIYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCARRGLL
10 LDYWGQGTTVTVSS **(hVH1, SEQ ID NO:3), or**
EVQLVESGGGLIQPGGSLRLSCAASGFTFSNYGMHWVRQAPGKGLEWVSYI
SSGSSTIYYADSVKGRFTISRDNSKNTLYLQMTSLRAEDTAVYYCARRGLL
LDYWGQGTTLTVSS **(hVH2, SEQ ID NO:4), or**
EVQLQESGGGVVQPGGSLRLSCAASGFTFSNYGMHWIRQAPGKGLEWVSYI
15 SSGSSTIYYADSVKGRFTISRDNSKNTLYLQMNSLRSEDTAVYYCARRGLL
LDYWGQGTLVTVSS **(hVH3, SEQ ID NO:5), or**
EVQLVESGGGLVQPGGSLRLSCSASGFTFSNYGMHWVRQAPGKGLEYVSYI
SSGSSTIYYADTVKGRFTISRDNSKNTLYLQMSSLRAEDTAVYYCVKRGLL
LDYWGQGTLVTVSS **(hVH4, SEQ ID NO:6), or**
20 EVQLVESGGGLVQPGGSLRLSCAASGFTFSNYGMHWVRQAPGKGLEWVSYI
SSSSSTIYYADSVKGRFTISRDNAKNSLYLQMNSLRAEDTAVYYCARRGLL
LDYWGQGTTVTVSS **(variants 2, 6 and 10, SEQ ID NO:46), or**
EVQLVESGGGVVQPGGSLRLSCAASGFTFSNYGMHWVRQAPEKGLEWVSYI
SSSSSTIYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCARRGLL
25 LDYWGQGTTVTVSS **(variants 3, 7 and 11, SEQ ID NO:47), or**
EVQLVESGGGDVKPGGSLRLSCAASGFTFSNYGMHWVRQAPEKGLEWVSYI
SSSSSTIYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCARRGLL
LDYWGQGTTVTVSS **(variants 4, 8 and 12, SEQ ID NO:48), or**
EVQLVESGGGLVQPGGSLRLSCAASGFTFSNYGMHWVRQAPGKGLEWVSYI
30 SSGSSTIYYADSVKGRFTISRDNAKNSLYLQMNSLRAEDTAVYYCARRGLL
LDYWGQGTTVTVSS **(variants 13, 16 and 19, SEQ ID NO:50), or**

CA 03149421 2022-2-24

EVQLVESGGGVVQPGGSLRLSCAASGFTFSNYGMHWVRQAPEKGLEWVSYI
SSGSSTIYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCARRGLL
LDYWGQGTTVTVSS (variants 14 and 17, SEQ ID NO:51), or
EVQLVESGGGDVKPGGSLRLSCAASGFTFSNYGMHWVRQAPEKGLEWVSYI
SSGSSTIYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCARRGLL
LDYWGQGTTVTVSS (variants 15 and 18, SEQ ID NO:52).

b. Humanized 3E10 Light Chain Variable Regions

In some embodiments, a humanized 3E10 light chain variable domain includes
DIQMTQSPSSLSASVGDRVTITCRASKSVSTSSYSYLAWYQQKPEKAPKLL
IKYASYLQSGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQHSREFPWTF
GAGTKLELK (hVL1, SEQ ID NO:9), or
DIQMTQSPSSLSASVGDRVTISCRASKSVSTSSYSYMHWYQQKPEKAPKLL
IKYASYLQSGVPSRFSGSGSGTDFTLTISSLQPEDVATYYCQHSREFPWTF
GAGTKLELK (hVL2, SEQ ID NO:10), or
DIVLTQSPASLAVSPGQRATITCRASKSVSTSSYSYMHWYQQKPGQPPKLL
IYYASYLESGVPARFSGSGSGTDFTLTINPVEANDTANYYCQHSREFPWTF
GQGTKVEIK (hVL3, SEQ ID NO:11)
DIQMTQSPSSLSASLGDRATITCRASKSVSTSSYSYMHWYQQKPGQPPKLL
IKYASYLESGVPSRFSGSGSGTDFTLTISSLQPEDAATYYCQHSREFPWTF
GGGTKVEIK (variants 2, 3 and 4, SEQ ID NO:53)
DIQMTQSPSSLSASLGDRATITCRASKSVSTSSYSYMHWYQQKPGQAPKLL
IKYASYLESGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQHSREFPWTF
GQGTKVEIK (variants 6, 7 and 8, SEQ ID NO:54)
DIQMTQSPSSLSASVGDRVTITCRASKSVSTSSYSYMHWYQQKPGKAPKLL
IKYASYLESGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQHSREFPWTF
GQGTKVEIK (variants 10, 11 and 12, SEQ ID NO:55)
DIQMTQSPSSLSASLGDRATITCRASKTVSTSSYSYMHWYQQKPGQPPKLL
IKYASYLESGVPSRFSGSGSGTDFTLTISSLQPEDAATYYCQHSREFPWTF
GGGTKVEIK (variants 13, 14 and 15, SEQ ID NO:56)

CA 03149421 2022-2-24

DIQMTQSPSSLSASVGDRVTITCRASKTVSTSSYSYMHWYQQKPGKAPKLL
IKYASYLESGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQHSREFPWTF
GQGTKVEIK (variants 16, 17 and 18, SEQ ID NO:57)
DIQMTQSPSSLSASLGDRATITCRASKTVSTSSYSYMHWYQQKPGQAPKLL
5 IKYASYLESGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQHSREFPWTF
GQGTKVEIK (variant 19, SEQ ID NO:58)

c. Cell Penetration and Nuclear Localization

The disclosed compositions and methods typically utilize antibodies that maintain the ability to penetrate cells, and optionally nuclei.

10 The mechanisms of cellular internalization by autoantibodies are diverse. Some are taken into cells through electrostatic interactions or FcR-mediated endocytosis, while others utilize mechanisms based on association with cell surface myosin or calreticulin, followed by endocytosis (Ying-Chyi et al., *Eur J Immunol* 38, 3178-3190 (2008), Yanase et al., *J Clin Invest* 100,
15 25-31 (1997)). 3E10 penetrates cells in an Fc-independent mechanism (as evidenced by the ability of 3E10 fragments lacking an Fc to penetrate cells) but involves presence of the nucleoside transporter ENT2 (Weisbart et al., *Sci Rep* 5:12022. doi: 10.1038/srep12022. (2015), Zack et al., *J Immunol* 157, 2082-2088 (1996), Hansen et al., *J Biol Chem* 282, 20790-20793
20 (2007)). Thus, in some embodiments, the antibodies utilized in the disclosed compositions and methods are ones that penetrates cells in an Fc-independent mechanism but involves presence of the nucleoside transporter ENT2.

Mutations in 3E10 that interfere with its ability to bind DNA may
25 render the antibody incapable of nuclear penetration. Thus, typically the disclosed variants and humanized forms of the antibody maintain the ability to bind nucleic acids, particularly DNA. In addition, 3E10 scFv has previously been shown capable of penetrating into living cells and nucleic in an ENT2-dependent manner, with efficiency of uptake impaired in ENT2-
30 deficient cells (Hansen, et al., *J. Biol. Chem.* 282, 20790-20793 (2007)). Thus, in some embodiments, the disclosed variants and humanized forms of the antibody maintain the ability penetrate into cell nuclei in an ENT-dependent, preferably ENT2-dependent manner.

As discussed in WO 2019/152806 and WO 2019/152808 some humanized 3E10 variant were found to penetrate cell nuclei more efficiently than the original murine 3E10 (D31N) di-scFv, while others were found to have lost the ability to penetrate nuclei. In particular, variants 10 and 13
5 penetrated nuclei very well compared to the murine antibody.

Potential bipartite nuclear localization signals (NLS) in humanized 3E10 VL have been identified and may include part or all of the following sequences:

RASKSVSTSSYSYMHWYQQKPGQPPKLLIKY (SEQ ID NO:88);
10 RASKTVSTSSYSYMHWYQQKPGQPPKLLIKY (SEQ ID NO:89); or
*RVTIT*CRASKSVSTSSYSYMHWYQQKPGKAPKL (SEQ ID NO:90).

An exemplary consensus NLS can be, or include, (X)RASKTVSTSSYSYMHWYQQKPGQPPKLL(X)KY (where (X) = any residue, but preferentially is a basic residue (R or K) (SEQ ID NO:91) or a
15 variant thereof with at least 60, 65, 70, 75, 80, 85, 90, 95, 96, 97, 98, 99 percent sequence identity to SEQ ID NO:53.

Thus, in some embodiments, particularly where nuclear importation is important, the disclosed antibodies may include the sequence of any one of SEQ ID NOS:88-91, or fragments and variants thereof (e.g., 70, 75, 80, 85,
20 90, 95, 96, 97, 98, 99, or 100% amino acid sequence identity with any one of SEQ ID NOS:88-91) that can translocate into the nucleus of a cell.

Presence of an NLS indicates that 3E10 may cross the nuclear envelope via the nuclear import pathway. In some embodiments, the NLS improves importation by interacting with one or more members of the import
25 pathway. Thus, in some embodiments, the NLS can bind to importin-β, an importin-β/importin-α heterodimer, or a combination thereof.

3. Nucleic Acid Binding

The disclosed compositions and methods typically utilize antibodies that maintain the ability to bind nucleic acids such as DNA, RNA, or a
30 combination thereof.

The Examples below illustrate molecular modeling of 3E10 and additional 3E10 variants. Molecular modeling of 3E10 (Pymol) revealed a putative Nucleic Acid Binding pocket (NAB1) (see, e.g., Figures 14A and 14B), and illustrated with underlining the sequences below.

CA 03149421 2022-2-24

WT HEAVY CHAIN scFv SEQUENCE

E VQLVESGGGL VKPGGSRKLS CAASGFTFSD YGMHWVRQAP EKGLEWVAYI SSGSSTIYYA DTVKGRFTIS RDNAKNTLFL QMTSLRSEDT AMYYCARRGL LLDYWGQGTT LTVS (SEQ ID NO:92)

5 ***LIGHT CHAIN scFv SEQUENCE***

D IVLTQSPASL AVSLGQRATI SCRASKSVST SSYSYMHWYQ QKPGQPPKLL IKYASYLESG VPARFSGSGS GTDFTLNIHP VEEEDAATYY CQHSREFPWT FGGGTKLEIK RADAAPGGGG SGGGGSGGGGS (SEQ ID NO:93)

In some embodiments, the disclosed antibodies include some or all of 10 the underlined NAB1 sequences. In some embodiments, the antibodies include a variant sequence that has an altered ability of bind nucleic acids. In some embodiments, the mutations (e.g., substitutions, insertions, and/or deletions) in the NAB1 improve binding of the antibody to nucleic acids such as DNA, RNA, or a combination thereof. In some embodiments, the 15 mutations are conservative substitutions. In some embodiments, the mutations increase the cationic charge of the NAB1 pocket.

As discussed and exemplified herein, mutation of aspartic acid at residue 31 of CDR1 to asparagine increased the cationic charge of this residue and enhanced nucleic acid binding and delivery *in vivo* (3E10- 20 D31N).

Additional exemplary variants include mutation of aspartic acid at residue 31 of CDR1 to arginine (3E10-D31R), which modeling indicates expands cationic charge, or lysine (3E10-D31K) which modeling indicates changes charge orientation. Thus, in some embodiments, the 3E10 binding 25 protein includes a D31R or D31K substitution.

All of the sequences disclosed herein having the residue corresponding to 3E10 D31 or N31, are expressly disclosed with a D31R or D31K or N31R or N31K substitution therein.

Molecular modeling of 3E10 (Pymol) revealed a putative Nucleic 30 Acid Binding pocket (NAB1) (Figures 14A-14B). Mutation of aspartic acid at residue 31 of CDR1 to asparagine increased the cationic charge of this residue and enhanced nucleic acid binding and delivery in vivo (3E10- D31N).

CA 03149421 2022-2-24

Mutation of aspartic acid at residue 31 of CDR1 to arginine (3E10-D31R), further expanded the cationic charge while mutation to lysine (3E10-D31K) changed charge orientation (Figure 14A).

NAB1 amino acids predicted from molecular modeling have been underlined in the heavy and light chain sequences above. Figure 14B is an illustration showing molecular modeling of 3E10-scFv (Pymol) with NAB1 amino acid residues illustrated with punctate dots.

4. Fragments, Variants, and Fusion Proteins

The anti-DNA antibody can be composed of an antibody fragment or fusion protein including an amino acid sequence of a variable heavy chain and/or variable light chain that is at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, or 100% identical to the amino acid sequence of the variable heavy chain and/or light chain of 3E10 or a humanized form thereof (e.g., any of SEQ ID NOS:1-11 or 46-58, or the heavy and/or light chains of any of SEQ ID NOS:12-14).

The anti-DNA antibody can be composed of an antibody fragment or fusion protein that includes one or more CDR(s) that is at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, or 100% identical to the amino acid sequence of the CDR(s) of 3E10, or a variant or humanized form thereof (e.g., CDR(s) of any of SEQ ID NOS:1-11 or 46-58, or SEQ ID NOS:12-14, or SEQ ID NOS:15-30 or 42-45). The determination of percent identity of two amino acid sequences can be determined by BLAST protein comparison. In some embodiments, the antibody includes one, two, three, four, five, or all six of the CDRs of the above-described preferred variable domains.

Preferably, the antibody includes one of each of a heavy chain CDR1, CDR2, and CDR3 in combination with one of each of a light chain CDR1, CDR2, and CDR3.

Predicted complementarity determining regions (CDRs) of the light chain variable sequence for 3E10 are provided above. See also GenBank: AAA65681.1 - *immunoglobulin light chain, partial [Mus musculus]* and GenBank: L34051.1 - *Mouse Ig rearranged kappa-chain mRNA V-region*.

CA 03149421 2022-2-24

Predicted complementarity determining regions (CDRs) of the heavy chain variable sequence for 3E10 are provide above. See also, for example, Zack, et al., *Immunology and Cell Biology*, 72:513-520 (1994), GenBank Accession number AAA65679.1. Zach, et al., *J. Immunol*. 154 (4), 1987-1994 (1995) and GenBank: L16982.1 - *Mouse Ig reagrranged H-chain gene, partial cds.*

Thus, in some embodiments, the cell-penetrating antibody contains the CDRs, or the entire heavy and light chain variable regions, of SEQ ID NO:1 or 2, or the heavy chain region of SEQ ID NO:12 or 13; or a humanized form thereof in combination with SEQ ID NO:7 or 8, or the light chain region of SEQ ID NO:14; or a humanized form thereof. In some embodiments, the cell-penetrating antibody contains the CDRs, or the entire heavy and light chain variable regions, of SEQ ID NO:3, 4, 5, or 6 in combination with SEQ ID NO:9, 10, or 11. In some embodiments, the cell-penetrating antibody contains the CDRs, or the entire heavy and light chain variable regions, of any one of SEQ ID NO:46-48 or 50-52 in combination with any one of SEQ ID NO:53-58.

All of the sequences disclosed herein having the residue corresponding to 3E10 D31 or N31, are expressly disclosed with a D31R or D31K or N31R or N31K substitution therein. Thus, in some embodiments, the 3E10 binding protein is a variant of any of the foregoing or following sequences wherein the amino acid residue corresponding to residue 31 of the 3E10 heavy chain is substituted with arginine (R) or lysine (K).

Also included are fragments of antibodies which have bioactivity. The fragments, whether attached to other sequences or not, include insertions, deletions, substitutions, or other selected modifications of particular regions or specific amino acids residues, provided the activity of the fragment is not significantly altered or impaired compared to the nonmodified antibody or antibody fragment.

Techniques can also be adapted for the production of single-chain antibodies specific to an antigenic protein of the present disclosure. Methods for the production of single-chain antibodies are well known to those of skill in the art. A single chain antibody can be created by fusing together the variable domains of the heavy and light chains using a short peptide linker,

CA 03149421 2022-2-24

thereby reconstituting an antigen binding site on a single molecule. Single-chain antibody variable fragments (scFvs) in which the C-terminus of one variable domain is tethered to the N-terminus of the other variable domain via a 15 to 25 amino acid peptide or linker have been developed without
5 significantly disrupting antigen binding or specificity of the binding. The linker is chosen to permit the heavy chain and light chain to bind together in their proper conformational orientation.

The anti-DNA antibodies can be modified to improve their therapeutic potential. For example, in some embodiments, the cell-
10 penetrating anti-DNA antibody is conjugated to another antibody specific for a second therapeutic target in the cytoplasm and/or nucleus of a target cell. For example, the cell-penetrating anti-DNA antibody can be a fusion protein containing 3E10 Fv and a single chain variable fragment of a monoclonal antibody that specifically binds the second therapeutic target. In other
15 embodiments, the cell-penetrating anti-DNA antibody is a bispecific antibody having a first heavy chain and a first light chain from 3E10 and a second heavy chain and a second light chain from a monoclonal antibody that specifically binds the second therapeutic target.

Bispecific antibodies and other binding proteins having a first heavy
20 chain and a first light chain from 3E10 and a second heavy chain and a second light chain from a monoclonal antibody that specifically binds a second target are discussed in Weisbart, et al., *Mol. Cancer Ther.*, 11(10):2169-73 (2012), and Weisbart, et al., *Int. J. Oncology*, 25:1113-8 (2004), and U.S. Patent Application No. 2013/0266570, which are
25 specifically incorporated by reference in their entireties. In some embodiments, the second target is specific for a target cell-type, tissue, organ etc. Thus the second heavy chain and second light chain can serve as a targeting moiety that targets the complex to the target cell-type, tissue, organ. In some embodiments, the second heavy chain and second light chain target,
30 hematopoietic stem cells, $CD34^{+}$ cells, T cells or any another preferred cell type, e.g., by targeting a receptor or ligand expressed on the preferred cell type. In some embodiments, the second heavy chain and second light chain target the thymus, spleen, or cancer cells.

CA 03149421 2022-2-24

In some embodiments, particularly those for targeting T cell in vivo, for example, for in vivo production of CAR T cells, immune cell or T cell markers such as CD3, CD7, or CD8 can be targeted. For example, anti-CD8 antibodies and anti-CD3 Fab fragments have both been used to target T cells
5 in vivo (Pfeiffer, et al., *EMBO Mol Med.*, 10(11) (2018). pii: e9158. doi: 10.15252/emmm.201809158., Smith, et al., *Nat Nanotechnol.*, 12(8):813-820 (2017). doi: 10.1038/nnano.2017.57). Thus, in some embodiments, the 3E10 antibody or antigen binding fragment or fusion protein is a bispecific antibody part of which can bind specifically to CD3, CD7, CD8, or another
10 immune cell (e.g., T cell) marker, or a marker for a specific tissue such as the thymus, spleen, or liver.

Divalent single-chain variable fragments (di-scFvs) can be engineered by linking two scFvs. This can be done by producing a single peptide chain with two VH and two VL regions, yielding tandem scFvs.
15 ScFvs can also be designed with linker peptides that are too short for the two variable regions to fold together (about five amino acids), forcing scFvs to dimerize. This type is known as diabodies. Diabodies have been shown to have dissociation constants up to 40-fold lower than corresponding scFvs, meaning that they have a much higher affinity to their target. Still shorter
20 linkers (one or two amino acids) lead to the formation of trimers (triabodies or tribodies). Tetrabodies have also been produced. They exhibit an even higher affinity to their targets than diabodies. In some embodiments, the anti-DNA antibody may contain two or more linked single chain variable fragments of 3E10 (*e.g.*, 3E10 di-scFv, 3E10 tri-scFv), or conservative
25 variants thereof. In some embodiments, the anti-DNA antibody is a diabody or triabody (*e.g.*, 3E10 diabody, 3E10 triabody). Sequences for single and two or more linked single chain variable fragments of 3E10 are provided in WO 2017/218825 and WO 2016/033321.

The function of the antibody may be enhanced by coupling the
30 antibody or a fragment thereof with a therapeutic agent. Such coupling of the antibody or fragment with the therapeutic agent can be achieved by making an immunoconjugate or by making a fusion protein, or by linking the antibody or fragment to a nucleic acid such as DNA or RNA (*e.g.*, siRNA), comprising the antibody or antibody fragment and the therapeutic agent.

CA 03149421 2022-2-24

A recombinant fusion protein is a protein created through genetic engineering of a fusion gene. This typically involves removing the stop codon from a cDNA sequence coding for the first protein, then appending the cDNA sequence of the second protein in frame through ligation or
5 overlap extension PCR. The DNA sequence will then be expressed by a cell as a single protein. The protein can be engineered to include the full sequence of both original proteins, or only a portion of either. If the two entities are proteins, often linker (or "spacer") peptides are also added which make it more likely that the proteins fold independently and behave as
10 expected.

In some embodiments, the cell-penetrating antibody is modified to alter its half-life. In some embodiments, it is desirable to increase the half-life of the antibody so that it is present in the circulation or at the site of treatment for longer periods of time. For example, it may be desirable to
15 maintain titers of the antibody in the circulation or in the location to be treated for extended periods of time. In other embodiments, the half-life of the anti-DNA antibody is decreased to reduce potential side effects. Antibody fragments, such as 3E10Fv may have a shorter half-life than full size antibodies. Other methods of altering half-life are known and can be
20 used in the described methods. For example, antibodies can be engineered with Fc variants that extend half-life, e.g., using Xtend™ antibody half-life prolongation technology (Xencor, Monrovia, CA).

a. Linkers

The term "linker" as used herein includes, without limitation, peptide
25 linkers. The peptide linker can be any size provided it does not interfere with the binding of the epitope by the variable regions. In some embodiments, the linker includes one or more glycine and/or serine amino acid residues. Monovalent single-chain antibody variable fragments (scFvs) in which the C-terminus of one variable domain are typically tethered to the
30 N-terminus of the other variable domain via a 15 to 25 amino acid peptide or linker. The linker is chosen to permit the heavy chain and light chain to bind together in their proper conformational orientation. Linkers in diabodies, triabodies, etc., typically include a shorter linker than that of a monovalent scFv as discussed above. Di-, tri-, and other multivalent scFvs typically

CA 03149421 2022-2-24

include three or more linkers. The linkers can be the same, or different, in length and/or amino acid composition. Therefore, the number of linkers, composition of the linker(s), and length of the linker(s) can be determined based on the desired valency of the scFv as is known in the art. The linker(s)
5 can allow for or drive formation of a di-, tri-, and other multivalent scFv.

For example, a linker can include 4-8 amino acids. In a particular embodiment, a linker includes the amino acid sequence GQSSRSS (SEQ ID NO:31). In another embodiment, a linker includes 15-20 amino acids, for example, 18 amino acids. In a particular embodiment, the linker includes the
10 amino acid sequence GQSSRSSSGGGSSGGGGS (SEQ ID NO:32). Other flexible linkers include, but are not limited to, the amino acid sequences Gly-Ser, Gly-Ser-Gly-Ser (SEQ ID NO:33), Ala-Ser, Gly-Gly-Gly-Ser (SEQ ID NO:34), $(Gly_4\text{-}Ser)_2$ (SEQ ID NO:35) and $(Gly_4\text{-}Ser)_4$ (SEQ ID NO:36), and $(Gly\text{-}Gly\text{-}Gly\text{-}Gly\text{-}Ser)_3$ (SEQ ID NO:37).

15 Other exemplary linkers include, for example,
RADAAPGGGGSGGGGSGGGGS (SEQ ID NO:59) and
ASTKGPSVFPLAPLESSGS (SEQ ID NO:60).

b. Exemplary anti-DNA scFv Sequences

Exemplary murine 3E10 scFv sequences, including mono-, di-, and
20 tri- scFv are disclosed in WO 2016/033321, WO 2017/218825, WO 2019/018426, and WO/2019/018428, and provided below. Cell-penetrating antibodies for use in the disclosed compositions and methods include exemplary scFv, and fragments and variants thereof.

The amino acid sequence for scFv 3E10 (D31N) is:

25 AGIHDIVLTQSPASLAVSLGQRATISCRASKSVSTSSYSYMHWYQQKPGQP
PKLLIKYASYLESGVPARFSGSGSGTDFTLNIHPVEEEDAATYYCQHSREF
PWTFGGGTKLEIKRADAAPGGGGSGGGGSGGGGSEVQLVESGGGLVKPGGS
RKLSCAASGFTFSNYGMHWVRQAPEKGLEWVAYISSGSSTIYYADTVKGRF
TISRDNAKNTLFLQMTSLRSEDTAMYYCARRGLLLDYWGQGTTLTVSSLEQ
30 KLISEEDLNSAVDHHHHHH
(SEQ ID NO:38).

CA 03149421 2022-2-24

Annotation of scFv Protein Domains with Reference to SEQ ID NO:38

- AGIH sequence increases solubility (amino acids 1-4 of SEQ ID NO:38)
- Vk variable region (amino acids 5-115 of SEQ ID NO:38)
- Initial (6 aa) of light chain CH1 (amino acids 116-121 of SEQ ID NO:38)
- $(GGGGS)_3$ (SEQ ID NO:37) linker (amino acids 122-136 of SEQ ID NO:38)
- VH variable region (amino acids 137-252 of SEQ ID NO:38)
- Myc tag (amino acids 253-268 SEQ ID NO:38)
- His 6 tag (amino acids 269-274 of SEQ ID NO:38)

Amino acid sequence of 3E10 di-scFv (D31N)

Di-scFv 3E10 (D31N) is a di-single chain variable fragment including 2X the heavy chain and light chain variable regions of 3E10 and wherein the aspartic acid at position 31 of the heavy chain is mutated to an asparagine. The amino acid sequence for di-scFv 3E10 (D31N) is:

AGIHDIVLTQSPASLAVSLGQRATISCRASKSVSTSSYSYMHWYQQKPGQP
PKLLIKYASYLESGVPARFSGSGSGTDFTLNIHPVEEEDAATYYCQHSREF
PWTFGGGTKLEIKRADAAPGGGGSGGGGSGGGGSEVQLVESGGGLVKPGGS
RKLSCAASGFTFSNYGMHWVRQAPEKGLEWVAYISSGSSTIYYADTVKGRF
TISRDNAKNTLFLQMTSLRSEDTAMYYCARRGLLLDYWGQGTTLTVSSAST
KGPSVFPLAPLESSGSDIVLTQSPASLAVSLGQRATISCRASKSVSTSSYS
YMHWYQQKPGQPPKLLIKYASYLESGVPARFSGSGSGTDFTLNIHPVEEED
AATYYCQHSREFPWTFGGGTKLEIKRADAAPGGGGSGGGGSGGGGSEVQLV
ESGGGLVKPGGSRKLSCAASGFTFSNYGMHWVRQAPEKGLEWVAYISSGSS
TIYYADTVKGRFTISRDNAKNTLFLQMTSLRSEDTAMYYCARRGLLLDYWG
QGTTLTVSSLEQKLISEEDLNSAVDHHHHHH
(SEQ ID NO:39).

Annotation of di-scFv Protein Domains with Reference to SEQ ID NO:39

- AGIH sequence increases solubility (amino acids 1-4 of SEQ ID NO:39)
- Vk variable region (amino acids 5-115 of SEQ ID NO:39)

CA 03149421 2022-2-24

- Initial (6 aa) of light chain CH1 (amino acids 116-121 of SEQ ID NO:39)
- $(GGGGS)_3$ (SEQ ID NO:37) linker (amino acids 122-136 of SEQ ID NO:39)
- VH variable region (amino acids 137-252 of SEQ ID NO:39)
- Linker between Fv fragments consisting of human IgG CH1 initial 13 amino acids (amino acids 253-265 of SEQ ID NO:39)
- Swivel sequence (amino acids 266-271 of SEQ ID NO:39)
- Vk variable region (amino acids 272-382 of SEQ ID NO:39)
- Initial (6 aa) of light chain CH1 (amino acids 383-388 of SEQ ID NO:39)
- $(GGGGS)_3$ (SEQ ID NO:37) linker (amino acids 389-403 of SEQ ID NO:39)
- VH variable region (amino acids 404-519 of SEQ ID NO:39)
- Myc tag (amino acids 520-535 of SEQ ID NO:39)
- His 6 tag (amino acids 536-541 of SEQ ID NO:39)

Amino acid sequence for tri-scFv

Tri-scFv 3E10 (D31N) is a tri-single chain variable fragment including 3X the heavy chain and light chain variable regions of 310E and wherein the aspartic acid at position 31 of the heavy chain is mutated to an asparagine. The amino acid sequence for tri-scFv 3E10 (D31N) is:

AGIHDIVLTQSPASLAVSLGQRATISCRASKSVSTSSYSYMHWYQQKPGQP
PKLLIKYASYLESGVPARFSGSGSGTDFTLNIHPVEEEDAATYYCQHSREF
PWTFGGGTKLEIKRADAAPGGGGSGGGGSGGGGSEVQLVESGGGLVKPGGS
RKLSCAASGFTFSNYGMHWVRQAPEKGLEWVAYISSGSSTIYYADTVKGRF
TISRDNAKNTLFLQMTSLRSEDTAMYYCARRGLLLDYWGQGTTLTVSSAST
KGPSVFPLAPLESSGSDIVLTQSPASLAVSLGQRATISCRASKSVSTSSYS
YMHWYQQKPGQPPKLLIKYASYLESGVPARFSGSGSGTDFTLNIHPVEEED
AATYYCQHSREFPWTFGGGTKLEIKRADAAPGGGGSGGGGSGGGGSEVQLV
ESGGGLVKPGGSRKLSCAASGFTFSNYGMHWVRQAPEKGLEWVAYISSGSS
TIYYADTVKGRFTISRDNAKNTLFLQMTSLRSEDTAMYYCARRGLLLDYWG
QGTTLTVSSASTKGPSVFPLAPLESSGSDIVLTQSPASLAVSLGQRATISC
RASKSVSTSSYSYMHWYQQKPGQPPKLLIKYASYLESGVPARFSGSGSGTD

CA 03149421 2022-2-24

FTLNIHPVEEEDAATYYCQHSREFPWTFGGGTKLEIKRADAAPGGGGSGGG
GSGGGGSEVQLVESGGGLVKPGGSRKLSCAASGFTFSNYGMHWVRQAPEKG
LEWVAYISSGSSTIYYADTVKGRFTISRDNAKNTLFLQMTSLRSEDTAMYY
CARRGLLLDYWGQGTTLTVSSLEQKLISEEDLNSAVDHHHHHH
5 (SEQ ID NO:40).

Annotation of tri-scFv Protein Domains with Reference to SEQ ID NO:40

- AGIH sequence increases solubility (amino acids 1-4 of SEQ ID NO:40)
- Vk variable region (amino acids 5-115 of SEQ ID NO:40)
- 10 Initial (6 aa) of light chain CH1 (amino acids 116-121 of SEQ ID NO:40)
- $(GGGGS)_3$ (SEQ ID NO:37) linker (amino acids 122-136 of SEQ ID NO:40)
- VH variable region (amino acids 137-252 of SEQ ID NO:40)
- 15 Linker between Fv fragments consisting of human IgG CH1 initial 13 amino acids (amino acids 253-265 of SEQ ID NO:40)
- Swivel sequence (amino acids 266-271 of SEQ ID NO:40)
- Vk variable region (amino acids 272-382 of SEQ ID NO:40)
- Initial (6 aa) of light chain CH1 (amino acids 383-388 of SEQ ID 20 NO:40)
- $(GGGGS)_3$ (SEQ ID NO:37) linker (amino acids 389-403 of SEQ ID NO:40)
- VH variable region (amino acids 404-519 of SEQ ID NO:40)
- Linker between Fv fragments consisting of human IgG C_H1 initial 13 25 amino acids (amino acids 520-532 of SEQ ID NO:40)
- Swivel sequence (amino acids 533-538 of SEQ ID NO:40)
- Vk variable region (amino acids 539-649 of SEQ ID NO:40)
- Initial (6 aa) of light chain CH1 (amino acids 650-655 of SEQ ID NO:40)
- 30 $(GGGGS)_3$ (SEQ ID NO:37) linker (amino acids 656-670 of SEQ ID NO:40)
- VH variable region (amino acids 671-786 of SEQ ID NO:40)
- Myc tag (amino acids 787-802 of SEQ ID NO:40)

CA 03149421 2022-2-24

- His 6 tag (amino acids 803-808 of SEQ ID NO:40)

WO 2016/033321 and Noble, et al., *Cancer Research*, 75(11):2285-2291 (2015), show that di-scFv and tri-scFv have some improved and additional activities compared to their monovalent counterpart. The subsequences corresponding to the different domains of each of the exemplary fusion proteins are also provided above. One of skill in the art will appreciate that the exemplary fusion proteins, or domains thereof, can be utilized to construct fusion proteins discussed in more detail above. For example, in some embodiments, the di-scFv includes a first scFv including a Vk variable region (e.g., amino acids 5-115 of SEQ ID NO:39, or a functional variant or fragment thereof), linked to a VH variable domain (e.g., amino acids 137-252 of SEQ ID NO:39, or a functional variant or fragment thereof), linked to a second scFv including a Vk variable region (e.g., amino acids 272-382 of SEQ ID NO:39, or a functional variant or fragment thereof), linked to a VH variable domain (e.g., amino acids 404-519 of SEQ ID NO:39, or a functional variant or fragment thereof). In some embodiments, a tri-scFv includes a di-scFv linked to a third scFv domain including a Vk variable region (e.g., amino acids 539-649 of SEQ ID NO:40, or a functional variant or fragment thereof), linked to a VH variable domain (e.g., amino acids 671-786 of SEQ ID NO:40, or a functional variant or fragment thereof).

The Vk variable regions can be linked to VH variable domains by, for example, a linker (e.g., $(GGGGS)_3$ (SEQ ID NO:37), alone or in combination with a (6 aa) of light chain CH1 (amino acids 116-121 of SEQ ID NO:39). Other suitable linkers are discussed above and known in the art. scFv can be linked by a linker (e.g., human IgG CH1 initial 13 amino acids (253-265) of SEQ ID NO:39), alone or in combination with a swivel sequence (e.g., amino acids 266-271 of SEQ ID NO:39). Other suitable linkers are discussed above and known in the art.

Therefore, a di-scFv can include amino acids 5-519 of SEQ ID NO:39. A tri-scFv can include amino acids 5-786 of SEQ ID NO:40. In some embodiments, the fusion proteins include additional domains. For example, in some embodiments, the fusion proteins include sequences that enhance solubility (e.g., amino acids 1-4 of SEQ ID NO:39). Therefore, in

CA 03149421 2022-2-24

some embodiments, a di-scFv can include amino acids 1-519 of SEQ ID NO:39. A tri-scFv can include amino acids 1-786 of SEQ ID NO:40. In some embodiments that fusion proteins include one or more domains that enhance purification, isolation, capture, identification, separation, etc., of the
5 fusion protein. Exemplary domains include, for example, Myc tag (e.g., amino acids 520-535 of SEQ ID NO:39) and/or a His tag (e.g., amino acids 536-541 of SEQ ID NO:39). Therefore, in some embodiments, a di-scFv can include the amino acid sequence of SEQ ID NO:39. A tri-scFv can include the amino acid sequence of SEQ ID NO:40. Other substitutable domains and
10 additional domains are discussed in more detail above.

An exemplary 3E10 humanized Fv sequence is discussed in WO 2016/033324:

DIVLTQSPASLAVSPGQRATITCRASKSVSTSSYSYMHWYQQKPGQPPKLL
IYYASYLESGVPARFSGSGSGTDFTLTINPVEANDTANYYCQHSREFPWTF
15 GQGTKVEIKGGGGSGGGGSGGGGSEVQLVESGGGLVQPGGSLRLSCSASGF
TFSNYGMHWVRQAPGKGLEYVSYISSGSSTIYYADTVKGRFTISRDNSKNT
LYLQMSSLRAEDTAVYYCVKRGLLLDYWGQGTLVTVSS (SEQ ID NO:41).

Exemplary 3E10 humanized di-scFv sequences are discussed in WO 2019/018426 and WO/2019/018428, and include:

20 DIQMTQSPSSLSASLGDRATITCRASKSVSTSSYSYMHWYQQKPGQPPKLL
IKYASYLESGVPSRFSGSGSGTDFTLTISSLQPEDAATYYCQHSREFPWTF
GGGTKVEIKRADAAPGGGGSGGGGSGGGGSEVQLVESGGGLVQPGGSLRLS
CAASGFTFSNYGMHWVRQAPGKGLEWVSYISSSSSTIYYADSVKGRFTISR
DNAKNSLYLQMNSLRAEDTAVYYCARRGLLLDYWGQGTTVTVSSASTKGPS
25 VFPLAPLESSGSDIQMTQSPSSLSASLGDRATITCRASKSVSTSSYSYMHW
YQQKPGQPPKLLIKYASYLESGVPSRFSGSGSGTDFTLTISSLQPEDAATY
YCQHSREFPWTFGGGTKVEIKRADAAPGGGGSGGGGSGGGGSEVQLVESGG
GLVQPGGSLRLSCAASGFTFSNYGMHWVRQAPGKGLEWVSYISSSSSTIYY
ADSVKGRFTISRDNAKNSLYLQMNSLRAEDTAVYYCARRGLLLDYWGQGTT
30 VTVSS (Variant 2, SEQ ID NO:61),

DIQMTQSPSSLSASLGDRATITCRASKSVSTSSYSYMHWYQQKPGQPPKLL
IKYASYLESGVPSRFSGSGSGTDFTLTISSLQPEDAATYYCQHSREFPWTF
GGGTKVEIKRADAAPGGGGSGGGGSGGGGSEVQLVESGGVVQPGGSLRLS
CAASGFTFSNYGMHWVRQAPEKGLEWVSYISSSSSTIYYADSVKGRFTISR

CA 03149421 2022-2-24

DNSKNTLYLQMNSLRAEDTAVYYCARRGLLLDYWGQGTTVTVSSASTKGPS
VFPLAPLESSGSDIQMTQSPSSLSASLGDRATITCRASKSVSTSSYSYMHW
YQQKPGQPPKLLIKYASYLESGVPSRFSGSGSGTDFTLTISSLQPEDAATY
YCQHSREFPWTFGGGTKVEIKRADAAPGGGGSGGGGSGGGGSEVQLVESGG
5 GVVQPGGSLRLSCAASGFTFSNYGMHWVRQAPEKGLEWVSYISSSSSTIYY
ADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCARRGLLLDYWGQGTT
VTVSS **(Variant 3, SEQ ID NO:62),**
DIQMTQSPSSLSASLGDRATITCRASKSVSTSSYSYMHWYQQKPGQPPKLL
IKYASYLESGVPSRFSGSGSGTDFTLTISSLQPEDAATYYCQHSREFPWTF
10 GGGTKVEIKRADAAPGGGGSGGGGSGGGGSEVQLVESGGGDVKPGGSLRLS
CAASGFTFSNYGMHWVRQAPEKGLEWVSYISSSSSTIYYADSVKGRFTISR
DNSKNTLYLQMNSLRAEDTAVYYCARRGLLLDYWGQGTTVTVSSASTKGPS
VFPLAPLESSGSDIQMTQSPSSLSASLGDRATITCRASKSVSTSSYSYMHW
YQQKPGQPPKLLIKYASYLESGVPSRFSGSGSGTDFTLTISSLQPEDAATY
15 YCQHSREFPWTFGGGTKVEIKRADAAPGGGGSGGGGSGGGGSEVQLVESGG
GDVKPGGSLRLSCAASGFTFSNYGMHWVRQAPEKGLEWVSYISSSSSTIYY
ADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCARRGLLLDYWGQGTT
VTVSS **(Variant 4, SEQ ID NO:63),**
DIQMTQSPSSLSASLGDRATITCRASKSVSTSSYSYMHWYQQKPGQAPKLL
20 IKYASYLESGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQHSREFPWTF
GQGTKVEIKRADAAPGGGGSGGGGSGGGGSEVQLVESGGGLVQPGGSLRLS
CAASGFTFSNYGMHWVRQAPGKGLEWVSYISSSSSTIYYADSVKGRFTISR
DNAKNSLYLQMNSLRAEDTAVYYCARRGLLLDYWGQGTTVTVSSASTKGPS
VFPLAPLESSGSDIQMTQSPSSLSASLGDRATITCRASKSVSTSSYSYMHW
25 YQQKPGQAPKLLIKYASYLESGVPSRFSGSGSGTDFTLTISSLQPEDFATY
YCQHSREFPWTFGQGTKVEIKRADAAPGGGGSGGGGSGGGGSEVQLVESGG
GLVQPGGSLRLSCAASGFTFSNYGMHWVRQAPGKGLEWVSYISSSSSTIYY
ADSVKGRFTISRDNAKNSLYLQMNSLRAEDTAVYYCARRGLLLDYWGQGTT
VTVSS **(Variant 6, SEQ ID NO:64),**
30 DIQMTQSPSSLSASLGDRATITCRASKSVSTSSYSYMHWYQQKPGQAPKLL
IKYASYLESGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQHSREFPWTF
GQGTKVEIKRADAAPGGGGSGGGGSGGGGSEVQLVESGGGVVQPGGSLRLS
CAASGFTFSNYGMHWVRQAPEKGLEWVSYISSSSSTIYYADSVKGRFTISR
DNSKNTLYLQMNSLRAEDTAVYYCARRGLLLDYWGQGTTVTVSSASTKGPS

CA 03149421 2022-2-24

VFPLAPLESSGSDIQMTQSPSSLSASLGDRATITCRASKSVSTSSYSYMHW
YQQKPGQAPKLLIKYASYLESGVPSRFSGSGSGTDFTLTISSLQPEDFATY
YCQHSREFPWTFGQGTKVEIKRADAAPGGGGSGGGGSGGGGSEVQLVESGG
GVVQPGGSLRLSCAASGFTFSNYGMHWVRQAPEKGLEWVSYISSSSSTIYY
5 ADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCARRGLLLDYWGQGTT
VTVSS **(Variant 7, SEQ ID NO:65),**
DIQMTQSPSSLSASLGDRATITCRASKSVSTSSYSYMHWYQQKPGQAPKLL
IKYASYLESGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQHSREFPWTF
GQGTKVEIKRADAAPGGGGSGGGGSGGGGSEVQLVESGGGDVKPGGSLRLS
10 CAASGFTFSNYGMHWVRQAPEKGLEWVSYISSSSSTIYYADSVKGRFTISR
DNSKNTLYLQMNSLRAEDTAVYYCARRGLLLDYWGQGTTVTVSSASTKGPS
VFPLAPLESSGSDIQMTQSPSSLSASLGDRATITCRASKSVSTSSYSYMHW
YQQKPGQAPKLLIKYASYLESGVPSRFSGSGSGTDFTLTISSLQPEDFATY
YCQHSREFPWTFGQGTKVEIKRADAAPGGGGSGGGGSGGGGSEVQLVESGG
15 GDVKPGGSLRLSCAASGFTFSNYGMHWVRQAPEKGLEWVSYISSSSSTIYY
ADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCARRGLLLDYWGQGTT
VTVSS **(Variant 8, SEQ ID NO:66),**
DIQMTQSPSSLSASVGDRVTITCRASKSVSTSSYSYMHWYQQKPGKAPKLL
IKYASYLESGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQHSREFPWTF
20 GQGTKVEIKRADAAPGGGGSGGGGSGGGGSEVQLVESGGGLVQPGGSLRLS
CAASGFTFSNYGMHWVRQAPGKGLEWVSYISSSSSTIYYADSVKGRFTISR
DNAKNSLYLQMNSLRAEDTAVYYCARRGLLLDYWGQGTTVTVSSASTKGPS
VFPLAPLESSGSDIQMTQSPSSLSASVGDRVTITCRASKSVSTSSYSYMHW
YQQKPGKAPKLLIKYASYLESGVPSRFSGSGSGTDFTLTISSLQPEDFATY
25 YCQHSREFPWTFGQGTKVEIKRADAAPGGGGSGGGGSGGGGSEVQLVESGG
GLVQPGGSLRLSCAASGFTFSNYGMHWVRQAPGKGLEWVSYISSSSSTIYY
ADSVKGRFTISRDNAKNSLYLQMNSLRAEDTAVYYCARRGLLLDYWGQGTT
VTVSS **(Variant 10, SEQ ID NO:67),**
DIQMTQSPSSLSASVGDRVTITCRASKSVSTSSYSYMHWYQQKPGKAPKLL
30 IKYASYLESGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQHSREFPWTF
GQGTKVEIKRADAAPGGGGSGGGGSGGGGSEVQLVESGGGVVQPGGSLRLS
CAASGFTFSNYGMHWVRQAPEKGLEWVSYISSSSSTIYYADSVKGRFTISR
DNSKNTLYLQMNSLRAEDTAVYYCARRGLLLDYWGQGTTVTVSSASTKGPS
VFPLAPLESSGSDIQMTQSPSSLSASVGDRVTITCRASKSVSTSSYSYMHW

CA 03149421 2022-2-24

YQQKPGKAPKLLIKYASYLESGVPSRFSGSGSGTDFTLTISSLQPEDFATY
YCQHSREFPWTFGQGTKVEIKRADAAPGGGGSGGGGSGGGGSEVQLVESGG
GVVQPGGSLRLSCAASGFTFSNYGMHWVRQAPEKGLEWVSYISSSSSTIYY
ADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCARRGLLLDYWGQGTT
5 VTVSS (Variant 11, SEQ ID NO:68),
DIQMTQSPSSLSASVGDRVTITCRASKSVSTSSYSYMHWYQQKPGKAPKLL
IKYASYLESGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQHSREFPWTF
GQGTKVEIKRADAAPGGGGSGGGGSGGGGSEVQLVESGGGDVKPGGSLRLS
CAASGFTFSNYGMHWVRQAPEKGLEWVSYISSSSSTIYYADSVKGRFTISR
10 DNSKNTLYLQMNSLRAEDTAVYYCARRGLLLDYWGQGTTVTVSSASTKGPS
VFPLAPLESSGSDIQMTQSPSSLSASVGDRVTITCRASKSVSTSSYSYMHW
YQQKPGKAPKLLIKYASYLESGVPSRFSGSGSGTDFTLTISSLQPEDFATY
YCQHSREFPWTFGQGTKVEIKRADAAPGGGGSGGGGSGGGGSEVQLVESGG
GDVKPGGSLRLSCAASGFTFSNYGMHWVRQAPEKGLEWVSYISSSSSTIYY
15 ADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCARRGLLLDYWGQGTT
VTVSS (Variant 12, SEQ ID NO:69),
DIQMTQSPSSLSASLGDRATITCRASKTVSTSSYSYMHWYQQKPGQPPKLLI
KYASYLESGVPSRFSGSGSGTDFTLTISSLQPEDAATYYCQHSREFPWTFGG
GTKVEIKRADAAPGGGGSGGGGSGGGGSEVQLVESGGGLVQPGGSLRLSCAA
20 SGFTFSNYGMHWVRQAPGKGLEWVSYISSGSSTIYYADSVKGRFTISRDNAK
NSLYLQMNSLRAEDTAVYYCARRGLLLDYWGQGTTVTVSSASTKGPSVFPLA
PLESSGSDIQMTQSPSSLSASLGDRATITCRASKTVSTSSYSYMHWYQQKPG
QPPKLLIKYASYLESGVPSRFSGSGSGTDFTLTISSLQPEDAATYYCQHSRE
FPWTFGGGTKVEIKRADAAPGGGGSGGGGSGGGGSEVQLVESGGGLVQPGGS
25 LRLSCAASGFTFSNYGMHWVRQAPGKGLEWVSYISSGSSTIYYADSVKGRFT
ISRDNAKNSLYLQMNSLRAEDTAVYYCARRGLLLDYWGQGTTVTVSS
(Variant 13, SEQ ID NO:70),
DIQMTQSPSSLSASLGDRATITCRASKTVSTSSYSYMHWYQQKPGQPPKLL
IKYASYLESGVPSRFSGSGSGTDFTLTISSLQPEDAATYYCQHSREFPWTF
30 GGGTKVEIKRADAAPGGGGSGGGGSGGGGSEVQLVESGGGVVQPGGSLRLS
CAASGFTFSNYGMHWVRQAPEKGLEWVSYISSGSSTIYYADSVKGRFTISR
DNSKNTLYLQMNSLRAEDTAVYYCARRGLLLDYWGQGTTVTVSSASTKGPS
VFPLAPLESSGSDIQMTQSPSSLSASLGDRATITCRASKTVSTSSYSYMHW
YQQKPGQPPKLLIKYASYLESGVPSRFSGSGSGTDFTLTISSLQPEDAATY

CA 03149421 2022-2-24

YCQHSREFPWTFGGGTKVEIKRADAAPGGGGSGGGGSGGGGSEVQLVESGG
GVVQPGGSLRLSCAASGFTFSNYGMHWVRQAPEKGLEWVSYISSGSSTIYY
ADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCARRGLLLDYWGQGTT
VTVSS (**Variant 14, SEQ ID NO:71**),
5 DIQMTQSPSSLSASLGDRATITCRASKTVSTSSYSYMHWYQQKPGQPPKLL
IKYASYLESGVPSRFSGSGSGTDFTLTISSLQPEDAATYYCQHSREFPWTF
GGGTKVEIKRADAAPGGGGSGGGGSGGGGSEVQLVESGGGDVKPGGSLRLS
CAASGFTFSNYGMHWVRQAPEKGLEWVSYISSGSSTIYYADSVKGRFTISR
DNSKNTLYLQMNSLRAEDTAVYYCARRGLLLDYWGQGTTVTVSSASTKGPS
10 VFPLAPLESSGSDIQMTQSPSSLSASLGDRATITCRASKTVSTSSYSYMHW
YQQKPGQPPKLLIKYASYLESGVPSRFSGSGSGTDFTLTISSLQPEDAATY
YCQHSREFPWTFGGGTKVEIKRADAAPGGGGSGGGGSGGGGSEVQLVESGG
GDVKPGGSLRLSCAASGFTFSNYGMHWVRQAPEKGLEWVSYISSGSSTIYY
ADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCARRGLLLDYWGQGTT
15 VTVSS (**Variant 15, SEQ ID NO:72**),
DIQMTQSPSSLSASVGDRVTITCRASKTVSTSSYSYMHWYQQKPGKAPKLL
IKYASYLESGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQHSREFPWTF
GQGTKVEIKRADAAPGGGGSGGGGSGGGGSEVQLVESGGGLVQPGGSLRLS
CAASGFTFSNYGMHWVRQAPGKGLEWVSYISSGSSTIYYADSVKGRFTISR
20 DNAKNSLYLQMNSLRAEDTAVYYCARRGLLLDYWGQGTTVTVSSASTKGPS
VFPLAPLESSGSDIQMTQSPSSLSASVGDRVTITCRASKTVSTSSYSYMHW
YQQKPGKAPKLLIKYASYLESGVPSRFSGSGSGTDFTLTISSLQPEDFATY
YCQHSREFPWTFGQGTKVEIKRADAAPGGGGSGGGGSGGGGSEVQLVESGG
GLVQPGGSLRLSCAASGFTFSNYGMHWVRQAPGKGLEWVSYISSGSSTIYY
25 ADSVKGRFTISRDNAKNSLYLQMNSLRAEDTAVYYCARRGLLLDYWGQGTT
VTVSS (**Variant 16, SEQ ID NO:73**),
DIQMTQSPSSLSASVGDRVTITCRASKTVSTSSYSYMHWYQQKPGKAPKLL
IKYASYLESGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQHSREFPWTF
GQGTKVEIKRADAAPGGGGSGGGGSGGGGSEVQLVESGGGVVQPGGSLRLS
30 CAASGFTFSNYGMHWVRQAPEKGLEWVSYISSGSSTIYYADSVKGRFTISR
DNSKNTLYLQMNSLRAEDTAVYYCARRGLLLDYWGQGTTVTVSSASTKGPS
VFPLAPLESSGSDIQMTQSPSSLSASVGDRVTITCRASKTVSTSSYSYMHW
YQQKPGKAPKLLIKYASYLESGVPSRFSGSGSGTDFTLTISSLQPEDFATY
YCQHSREFPWTFGQGTKVEIKRADAAPGGGGSGGGGSGGGGSEVQLVESGG

GVVQPGGSLRLSCAASGFTFSNYGMHWVRQAPEKGLEWVSYISSGSSTIYY
ADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCARRGLLLDYWGQGTT
VTVSS (Variant 17, SEQ ID NO:74),
DIQMTQSPSSLSASVGDRVTITCRASKTVSTSSYSYMHWYQQKPGKAPKLL
5 IKYASYLESGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQHSREFPWTF
GQGTKVEIKRADAAPGGGGSGGGGSGGGGSEVQLVESGGGDVKPGGSLRLS
CAASGFTFSNYGMHWVRQAPEKGLEWVSYISSGSSTIYYADSVKGRFTISR
DNSKNTLYLQMNSLRAEDTAVYYCARRGLLLDYWGQGTTVTVSSASTKGPS
VFPLAPLESSGSDIQMTQSPSSLSASVGDRVTITCRASKTVSTSSYSYMHW
10 YQQKPGKAPKLLIKYASYLESGVPSRFSGSGSGTDFTLTISSLQPEDFATY
YCQHSREFPWTFGQGTKVEIKRADAAPGGGGSGGGGSGGGGSEVQLVESGG
GDVKPGGSLRLSCAASGFTFSNYGMHWVRQAPEKGLEWVSYISSGSSTIYY
ADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCARRGLLLDYWGQGTT
VTVSS (Variant 18, SEQ ID NO:75), and
15 DIQMTQSPSSLSASLGDRATITCRASKTVSTSSYSYMHWYQQKPGQAPKLL
IKYASYLESGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQHSREFPWTF
GQGTKVEIKRADAAPGGGGSGGGGSGGGGSEVQLVESGGGLVQPGGSLRLS
CAASGFTFSNYGMHWVRQAPGKGLEWVSYISSGSSTIYYADSVKGRFTISR
DNAKNSLYLQMNSLRAEDTAVYYCARRGLLLDYWGQGTTVTVSSASTKGPS
20 VFPLAPLESSGSDIQMTQSPSSLSASLGDRATITCRASKTVSTSSYSYMHW
YQQKPGQAPKLLIKYASYLESGVPSRFSGSGSGTDFTLTISSLQPEDFATY
YCQHSREFPWTFGQGTKVEIKRADAAPGGGGSGGGGSGGGGSEVQLVESGG
GLVQPGGSLRLSCAASGFTFSNYGMHWVRQAPGKGLEWVSYISSGSSTIYY
ADSVKGRFTISRDNAKNSLYLQMNSLRAEDTAVYYCARRGLLLDYWGQGTT
25 VTVSS (Variant 19, SEQ ID NO:76).

c. Additional Sequences

Additional sequences that may used in the construction of anti-DNA antigen binding proteins, antibodies, fragments and fusion proteins include, but are not limited to,

30 EVQLVESGGGLVQPGGSLRLSCAASGFTFSNYGMHWVRQAPGKGLEWVSYI
SSGSSTIYYADSVKGRFTISRDNAKNSLYLQMNSLRAEDTAVYYCARRGLL
LDYWGQGTTVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPV
TVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHK
PSNTKVDKKVEPKSCDKTHTCPPCPAPEAAGGPSVFLFPPKPKDTLMISRT

CA 03149421 2022-2-24

PEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLT
VLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDEL
TKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSK
LTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK **(IgG1**
5 **L2345A/L235A heavy chain full length sequence, SEQ ID NO:77),**
ASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVH
TFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKV
(IgG1 constant heavy region 1, SEQ ID NO:78),
EPKSCDKTHTCP **(IgG1 hinge region, SEQ ID NO:79),**
10 PCPAPEAAGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWY
VDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALP
APIEKTISKAK **(IgG1 L2345A/L235A constant heavy region 2, SEQ ID NO:80),**
GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNY
15 KTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLS
LSPGK **(IgG1 constant heavy region 3, SEQ ID NO:81),**
EVQLVESGGGLVQPGGSLRLSCAASGFTFSNYGMHWVRQAPGKGLEWVSYI
SSGSSTIYYADSVKGRFTISRDNAKNSLYLQMNSLRAEDTAVYYCARRGLL
LDYWGQGTTVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPV
20 TVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHK
PSNTKVDKKVEPKSCDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRT
PEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYDSTYRVVSVLT
VLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDEL
TKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSK
25 LTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK **(IgG1 N297D heavy chain full length sequence, SEQ ID NO:82),**
PCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWY
VDGVEVHNAKTKPREEQYDSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALP
APIEKTISKAK **(IgG1 N297D constant heavy region 2, SEQ ID NO:83),**
30 EVQLVESGGGLVQPGGSLRLSCAASGFTFSNYGMHWVRQAPGKGLEWVSYI
SSGSSTIYYADSVKGRFTISRDNAKNSLYLQMNSLRAEDTAVYYCARRGLL
LDYWGQGTTVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPV
TVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHK
PSNTKVDKKVEPKSCDKTHTCPPCPAPEAAGGPSVFLFPPKPKDTLMISRT

PEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYDSTYRVVSVLT
VLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDEL
TKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSK
LTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK (IgG1
5 L2345A/L235A/N297D heavy chain full length sequence, SEQ ID NO:84),
PCPAPEAAGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWY
VDGVEVHNAKTKPREEQYDSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALP
APIEKTISKAK (IgG1 L2345A/L235A/N297D constant heavy region 2,
SEQ ID NO:85),
10 PCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWY
VDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALP
APIEKTISKAK (SEQ ID NO:86, Unmodified constant heavy region 2), and
DIQMTQSPSSLSASLGDRATITCRASKTVSTSSYSYMHWYQQKPGQPPKLL
IKYASYLESGVPSRFSGSGSGTDFTLTISSLQPEDAATYYCQHSREFPWTF
15 GGGTKLEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWK
VDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQG
LSSPVTKSFNRGEC (Light chain full length sequence, SEQ ID NO:87).

B. Cargo

As used in the methods provided herein, the 3E10 is typically
20 contacted with cells in complex with a nucleic acid cargo. The interaction between the antibody or binding protein and the nucleic acid cargo is non-covalent.

The nucleic acid cargo can be single stranded or double stranded. The nucleic acid cargo can be or include DNA, RNA, nucleic acid analogs,
25 or a combination thereof. As discussed in more detail below, nucleic acid analogs can be modified at the base moiety, sugar moiety, or phosphate backbone. Such modification can improve, for example, stability, hybridization, or solubility of the nucleic acid.

The nucleic acid cargo is typically functional in the sense that is or
30 encodes an agent that is biologically active once delivered into cells. Exemplary cargo is discussed in more detail below, but includes, for example, mRNA or DNA encoding polypeptides of interest including, for example expression constructs and vectors, inhibitory nucleic acids such as

CA 03149421 2022-2-24

siRNA, or nucleic acid encoding the inhibitory nucleic acid including, for example expression constructs and vectors.

The disclosed compositions can include a plurality of a single nucleic acid cargo molecule. In some embodiments, the compositions include a plurality of a multiplicity (e.g., 2, 3, 4, 5, 6, 7, 8, 9 10, or more) of different nucleic acid molecules.

In some embodiments, the cargo molecules are 0.001, 0.01, 1, 10's 100's, 1,000's, 10,000's, and/or 100,000's of kilobases in length.

In some embodiments, e.g., the cargo may be between 0.001 kb and 100 kb, or between 0.001 kb kb and 50 kb, or between 0.001 kb kb and 25 kb, or between 0.001 kb and 12.5 kb, or between 0.001 kb and 10 kb, or between 0.001 kb and 8 kb, or 0.001 kb and 5 kb, or between 0.001 kb and 2.5 kb, or between 0.001 kb and 1 kb, or between 0.01 kb and 100 kb, or between 0.01 kb kb and 50 kb, or between 0.01 kb kb and 25 kb, or between 0.01 kb and 12.5 kb, or between 0.01 kb and 10 kb, or between 0.01 kb and 8 kb, or 0.01 kb and 5 kb, or between 0.01 kb and 2.5 kb, or between 0.01 kb and 1 kb, or between 0.1 kb and 100 kb, or between 0.1 kb kb and 50 kb, or between 0.1 kb kb and 25 kb, or between 0.1 kb and 12.5 kb, or between 0.1 kb and 10 kb, or between 0.1 kb and 8 kb, or 0.1 kb and 5 kb, or between 0.1 kb and 2.5 kb, or between 0.1 kb and 1 kb, or between 1 kb and 100 kb, or between 1 kb kb and 50 kb, or between 1 kb kb and 25 kb, or between 1 kb and 12.5 kb, or between 1 kb and 10 kb, or between 1 kb and 8 kb, or 1 kb and 5 kb, or between 1 kb and 2.5 kb, each inclusive.

In some embodiments, e.g., the cargo may be between 0.2 kb and 10 kb, or between 0.2 kb and 5 kb, or between 0.2 kb and 2.5 kb, or between 0.2 kb and 1 kb, or between 0.2 kb and 0.5 kb, or between 0.2 kb and 0.25 kb, or between 0.5 kb and 10 kb, or between 0.5 kb and 5 kb, or between 1 kb and 5 kb, or between 1 kb and 3 kb, or between 2 kb and 10 kb, or between 3 kb and 5 kb.

It will be appreciated that for specific application the nucleic acid cargo may be one or more discrete lengths that, for example, falls within one of the foregoing ranges (inclusive), the specific values for each are expressly disclosed. For example, the size can be as small as a single nucleotide or nucleobase. In an exemplary application the cargo is a cyclic dinucleotide

CA 03149421 2022-2-24

like cGAMP, which is a STING agonist. In other embodiments, the cargo is a short oligomer. For example, oligomers as short as 8-mers can be used for anti-sense or splice switching. Slightly longer ones (e.g., 18 to 20 mers) can be used for gene editing.

5 **1. Forms of the Cargo**

The nucleic acid cargo is a nucleic acid and can be an isolated nucleic acid composition. As used herein, "isolated nucleic acid" refers to a nucleic acid that is separated from other nucleic acid molecules that are present in a mammalian genome, including nucleic acids that normally flank one or both
10 sides of the nucleic acid in a mammalian genome. The term "isolated" as used herein with respect to nucleic acids also includes the combination with any non-naturally-occurring nucleic acid sequence, since such non-naturally-occurring sequences are not found in nature and do not have immediately contiguous sequences in a naturally-occurring genome.

15 An isolated nucleic acid can be, for example, a DNA molecule, provided one of the nucleic acid sequences normally found immediately flanking that DNA molecule in a naturally-occurring genome is removed or absent. Thus, an isolated nucleic acid includes, without limitation, a DNA molecule that exists as a separate molecule independent of other sequences
20 (e.g., a chemically synthesized nucleic acid, or a cDNA or genomic DNA fragment produced by PCR or restriction endonuclease treatment), as well as recombinant DNA that is incorporated into a vector, an autonomously replicating plasmid, a virus (e.g., a retrovirus, lentivirus, adenovirus, or herpes virus), or into the genomic DNA of a prokaryote or eukaryote. In
25 addition, an isolated nucleic acid can include an engineered nucleic acid such as a recombinant DNA molecule that is part of a hybrid or fusion nucleic acid. A nucleic acid existing among hundreds to millions of other nucleic acids within, for example, a cDNA library or a genomic library, or a gel slice containing a genomic DNA restriction digest, is not to be considered an
30 isolated nucleic acid.

The nucleic acid sequences encoding polypeptides include genomic sequences. Also disclosed are mRNA/cDNA sequence wherein the exons have been deleted. Other nucleic acid sequences encoding polypeptides, such polypeptides that include the above-identified amino acid sequences

and fragments and variants thereof, are also disclosed. Nucleic acids encoding polypeptides may be optimized for expression in the expression host of choice. Codons may be substituted with alternative codons encoding the same amino acid to account for differences in codon usage between the organism from which the nucleic acid sequence is derived and the expression host. In this manner, the nucleic acids may be synthesized using expression host-preferred codons.

Nucleic acids can be in sense or antisense orientation, or can be, for example, complementary to a reference sequence encoding a polypeptide.

a. Vectors

The cargo can be a vector, for example, a vector encoding a polypeptide(s) and/or functional nucleic acid(s). Nucleic acids, such as those described above, can be inserted into vectors for expression in cells. As used herein, a "vector" is a replicon, such as a plasmid, phage, virus or cosmid, into which another DNA segment may be inserted so as to bring about the replication of the inserted segment. Vectors can be expression vectors. An "expression vector" is a vector that includes one or more expression control sequences, and an "expression control sequence" is a DNA sequence that controls and regulates the transcription and/or translation of another DNA sequence.

Nucleic acids in vectors can be operably linked to one or more expression control sequences. For example, the control sequence can be incorporated into a genetic construct so that expression control sequences effectively control expression of a coding sequence of interest. Examples of expression control sequences include promoters, enhancers, and transcription terminating regions. A promoter is an expression control sequence composed of a region of a DNA molecule, typically within 100 nucleotides upstream of the point at which transcription starts (generally near the initiation site for RNA polymerase II). To bring a coding sequence under the control of a promoter, it is necessary to position the translation initiation site of the translational reading frame of the polypeptide between one and about fifty nucleotides downstream of the promoter. Enhancers provide expression specificity in terms of time, location, and level. Unlike promoters, enhancers can function when located at various distances from the transcription site.

CA 03149421 2022-2-24

An enhancer also can be located downstream from the transcription initiation site. A coding sequence is "operably linked" and "under the control" of expression control sequences in a cell when RNA polymerase is able to transcribe the coding sequence into mRNA, which then can be translated into
5 the protein encoded by the coding sequence.

Suitable expression vectors include, without limitation, plasmids, cosmids, and viral vectors derived from, for example, bacteriophage, baculoviruses, tobacco mosaic virus, herpes viruses, cytomegalo virus, retroviruses, vaccinia viruses, adenoviruses, and adeno-associated viruses.
10 Numerous vectors and expression systems are commercially available from such corporations as Novagen (Madison, WI), Clontech (Palo Alto, CA), Stratagene (La Jolla, CA), and Invitrogen Life Technologies (Carlsbad, CA).

In some embodiments, the cargo is delivered into the cell and remains extrachromosomal. In some embodiments, the cargo is introduced into a
15 host cell and is integrated into the host cell's genome. As discussed in more detail below, the compositions can be used in methods of gene therapy. Methods of gene therapy can include the introduction into the cell of a polynucleotide that alters the genotype of the cell. Introduction of the polynucleotide can correct, replace, or otherwise alter the endogenous gene
20 via genetic recombination. Methods can include introduction of an entire replacement copy of a defective gene, a heterologous gene, or a small nucleic acid molecule such as an oligonucleotide. For example, a corrective gene can be introduced into a non-specific location within the host's genome.

In some embodiments, the cargo is a vector. Methods to construct
25 expression vectors containing genetic sequences and appropriate transcriptional and translational control elements are well known in the art. These methods include *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination. Expression vectors generally contain regulatory sequences and necessary elements for the translation
30 and/or transcription of the inserted coding sequence, which can be, for example, the polynucleotide of interest. The coding sequence can be operably linked to a promoter and/or enhancer to help control the expression of the desired gene product. Promoters used in biotechnology are of different types according to the intended type of control of gene expression. They can

CA 03149421 2022-2-24

be generally divided into constitutive promoters, tissue-specific or development-stage-specific promoters, inducible promoters, and synthetic promoters.

For example, in some embodiments, a polynucleotide of interest is operably linked to a promoter or other regulatory elements known in the art. Thus, the cargo can be a vector such as an expression vector. The engineering of polynucleotides for expression in a prokaryotic or eukaryotic system may be performed by techniques generally known to those of skill in recombinant expression. An expression vector typically includes one of the disclosed compositions under the control of one or more promoters. To bring a coding sequence "under the control of" a promoter, one positions the 5' end of the translational initiation site of the reading frame generally between about 1 and 50 nucleotides "downstream" of (i.e., 3' of) the chosen promoter. The "upstream" promoter stimulates transcription of the inserted DNA and promotes expression of the encoded recombinant protein or functional nucleic acid. This is the meaning of "recombinant expression" in the context used here.

Many standard techniques are available to construct expression vectors containing the appropriate nucleic acids and transcriptional/translational control sequences in order to achieve protein or peptide or functional nucleic acid expression in a variety of host-expression systems.

Expression vectors for use in mammalian cells ordinarily include an origin of replication (as necessary), a promoter located in front of the gene to be expressed, along with any necessary ribosome binding sites, RNA splice sites, polyadenylation site, and transcriptional terminator sequences. The origin of replication may be provided either by construction of the vector to include an exogenous origin, such as may be derived from SV40 or other viral (e.g., Polyoma, Adeno, VSV, BPV) source, or may be provided by the host cell chromosomal replication mechanism. If the vector is integrated into the host cell chromosome, the latter is often sufficient.

The promoters may be derived from the genome of mammalian cells (e.g., metallothionein promoter) or from mammalian viruses (e.g., the adenovirus late promoter; the vaccinia virus 7.5K promoter). Further, it is

CA 03149421 2022-2-24

also possible, and may be desirable, to utilize promoter or control sequences normally associated with the desired gene sequence, provided such control sequences are compatible with the host cell systems.

A number of viral based expression systems may be utilized, for
5 example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40 (SV40). The early and late promoters of SV40 virus are useful because both are obtained easily from the virus as a fragment which also contains the SV40 viral origin of replication. Smaller or larger SV40 fragments may also be used, provided there is
10 included the approximately 250 bp sequence extending from the HindIII site toward the BglI site located in the viral origin of replication.

In cases where an adenovirus is used as an expression vector, the coding sequences may be ligated to an adenovirus transcription/translation control complex, e.g., the late promoter and tripartite leader sequence. This
15 chimeric gene may then be inserted in the adenovirus genome by in vitro or in vivo recombination. Insertion in a non-essential region of the viral genome (e.g., region E1 or E3) will result in a recombinant virus that is viable and capable of expressing proteins in infected hosts.

Specific initiation signals may also be required for efficient
20 translation of the disclosed compositions. These signals include the ATG initiation codon and adjacent sequences. Exogenous translational control signals, including the ATG initiation codon, may additionally need to be provided. One of ordinary skill in the art would readily be capable of determining this need and providing the necessary signals. It is well known
25 that the initiation codon must be in-frame (or in-phase) with the reading frame of the desired coding sequence to ensure translation of the entire insert. These exogenous translational control signals and initiation codons can be of a variety of origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of appropriate transcription
30 enhancer elements or transcription terminators.

In eukaryotic expression, one will also typically desire to incorporate into the transcriptional unit an appropriate polyadenylation site if one was not contained within the original cloned segment. Typically, the poly A

addition site is placed about 30 to 2000 nucleotides "downstream" of the termination site of the protein at a position prior to transcription termination.

For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell lines that stably express constructs encoding proteins may be engineered. Rather than using expression vectors that contain viral origins of replication, host cells can be transformed with vectors controlled by appropriate expression control elements (e.g., promoter, enhancer, sequences, transcription terminators, polyadenylation sites, etc.), and a selectable marker. Following the introduction of foreign DNA, engineered cells may be allowed to grow for 1-2 days in an enriched medium, and then are switched to a selective medium. The selectable marker in the recombinant plasmid confers resistance to the selection and allows cells to stably integrate the plasmid into their chromosomes and grow to form foci, which in turn can be cloned and expanded into cell lines.

b. mRNAs

The cargo can be mRNA.

Chemical structures with the ability to promote stability and/or translation efficiency may also be used. For example, the RNA can have 5' and 3' UTRs. The length of the 3' UTR can, for example, exceed 100 nucleotides. In some embodiments the 3' UTR sequence is between 100 and 5000 nucleotides. In some embodiments, the 5' UTR is between zero and 3000 nucleotides in length. The length of 5' and 3' UTR sequences to be added to the coding region can be altered by different methods, including, but not limited to, designing primers for PCR that anneal to different regions of the UTRs. Using this approach, one of ordinary skill in the art can modify the 5' and 3' UTR lengths required to achieve optimal translation efficiency following delivery of the transcribed RNA.

The 5' and 3' UTRs can be the naturally occurring, endogenous 5' and 3' UTRs for the gene of interest. Alternatively, UTR sequences that are not endogenous to the gene of interest can be added by incorporating the UTR sequences into the forward and reverse primers or by any other modifications of the template. The use of UTR sequences that are not endogenous to the gene of interest can be useful for modifying the stability and/or translation efficiency of the RNA. For example, it is known that AU-

CA 03149421 2022-2-24

rich elements in 3' UTR sequences can decrease the stability of mRNA. Therefore, 3' UTRs can be selected or designed to increase the stability of the transcribed RNA based on properties of UTRs that are well known in the art.

5 In some embodiments, the 5' UTR contains the Kozak sequence of the endogenous gene. Alternatively, when a 5' UTR that is not endogenous to the gene of interest is being added by PCR as described above, a consensus Kozak sequence can be redesigned by adding the 5' UTR sequence. Kozak sequences can increase the efficiency of translation of 10 some RNA transcripts, but does not appear to be required for all RNAs to enable efficient translation. The requirement for Kozak sequences for many mRNAs is known in the art. In other embodiments the 5' UTR can be derived from an RNA virus whose RNA genome is stable in cells. In other embodiments various nucleotide analogues can be used in the 3' or 5' UTR 15 to impede exonuclease degradation of the mRNA.

In some embodiments, the mRNA has a cap on the 5' end, a 3' poly(A) tail, or a combination thereof which determine ribosome binding, initiation of translation and stability mRNA in the cell.

5'caps provide stability to RNA molecules. The 5' cap may, for 20 example, be $m^7G(5')ppp(5')G$, $m^7G(5')ppp(5')A$, G(5')ppp(5')G or G(5')ppp(5')A cap analogs, which are all commercially available. The 5' cap can also be an anti-reverse-cap-analog (ARCA) (Stepinski, et al., *RNA*, 7:1468-95 (2001)) or any other suitable analog. The 5' cap can be incorporated using techniques known in the art (Cougot, et al., *Trends in* 25 *Biochem. Sci.*, 29:436-444 (2001); Stepinski, et al., *RNA*, 7:1468-95 (2001); Elango, et al., *Biochim. Biophys. Res. Commun.*, 330:958-966 (2005)).

The RNAs can also contain an internal ribosome entry site (IRES) sequence. The IRES sequence may be any viral, chromosomal or artificially designed sequence which initiates cap-independent ribosome binding to 30 mRNA and facilitates the initiation of translation.

Generally, the length of a poly(A) tail positively correlates with the stability of the transcribed RNA. In one embodiment, the poly(A) tail is between 100 and 5000 adenosines.

CA 03149421 2022-2-24

A polyA segment can be produced during PCR by using a reverse primer containing a polyT tail, such as 100T tail (size can be, e.g., 50-5000 T), or after PCR by any other method, including, but not limited to, DNA ligation or *in vitro* recombination. Poly(A) tails also provide stability to RNAs and reduce their degradation. Poly(A) tails of RNAs can additionally or alternatively be extended following *in vitro* transcription with the use of a poly(A) polymerase, such as *E. coli* polyA polymerase (E-PAP).

Additionally, the attachment of different chemical groups to the 3' end can increase mRNA stability. Such attachment can contain modified/artificial nucleotides, aptamers and other compounds. For example, ATP analogs can be incorporated into the poly(A) tail using poly(A) polymerase. ATP analogs can further increase the stability of the RNA. Suitable ATP analogs include, but are not limited to, cordiocipin and 8-azaadenosine.

2. Sequence of the Cargo

a. Polypeptide of Interest

The cargo can encode one or more proteins. The cargo can be a polynucleotide that can be monocistronic or polycistronic. In some embodiments, polynucleotide is multigenic. The polynucleotide can be, for example, an mRNA or a expression construct such as a vector.

The cargo can encode one or more polypeptides of interest. The polypeptide can be any polypeptide. For example, the polypeptide encoded by the polynucleotide can be a polypeptide that provides a therapeutic or prophylactic effect to an organism or that can be used to diagnose a disease or disorder in an organism. For example, for treatment of cancer, autoimmune disorders, parasitic, viral, bacterial, fungal or other infections, the polynucleotide(s) to be expressed may encode a polypeptide that functions as a ligand or receptor for cells of the immune system, or can function to stimulate or inhibit the immune system of an organism.

In some embodiments, the polynucleotide supplements or replaces a polynucleotide that is defective in the organism.

In particular embodiments, the polynucleotide encodes dystrophin, utrophin, or a combination thereof. Such compositions may be administered

CA 03149421 2022-2-24

in an effective amount to treat a subject from a dystrophy, particularly a muscular dystrophy, for example, Duchenne's muscular dystrophy.

In another particular embodiment, the polynucleotide encodes antigen, e.g., an antigen that can be utilized in a vaccine formulation and associated methods. In a particular embodiment, polynucleotide encodes a viral antigen(s), for example, a SARS-CoV-2 antigen(s). Thus, compositions and methods of use thereof for protection against, and the treatment of, SARS-CoV-2 virus and viral infections and disease associate therewith including COVID19 are provided.

In some embodiments, the polynucleotide includes a selectable marker, for example, a selectable marker that is effective in a eukaryotic cell, such as a drug resistance selection marker. This selectable marker gene can encode a factor necessary for the survival or growth of transformed host cells grown in a selective culture medium. Typical selection genes encode proteins that confer resistance to antibiotics or other toxins, e.g., ampicillin, neomycin, methotrexate, kanamycin, gentamycin, Zeocin, or tetracycline, complement auxotrophic deficiencies, or supply critical nutrients withheld from the media.

In some embodiments, the polynucleotide includes a reporter gene. Reporter genes are typically genes that are not present or expressed in the host cell. The reporter gene typically encodes a protein which provides for some phenotypic change or enzymatic property. Examples of such genes are provided in Weising *et al. Ann. Rev. Genetics*, 22, 421 (1988). Preferred reporter genes include glucuronidase (GUS) gene and GFP genes.

b. Functional Nucleic Acids

The cargo can be or encode a functional nucleic acid. Functional nucleic acids are nucleic acid molecules that have a specific function, such as binding a target molecule or catalyzing a specific reaction. As discussed in more detail below, functional nucleic acid molecules can be divided into the following non-limiting categories: antisense molecules, siRNA, miRNA, aptamers, ribozymes, RNAi, and external guide sequences, and cyclic dinucleotides. The functional nucleic acid molecules can act as effectors, inhibitors, modulators, and stimulators of a specific activity possessed by a

CA 03149421 2022-2-24

target molecule, or the functional nucleic acid molecules can possess a *de novo* activity independent of any other molecules.

Functional nucleic acid molecules can interact with any macromolecule, such as DNA, RNA, polypeptides, or carbohydrate chains. Thus, functional nucleic acids can interact with the mRNA or the genomic DNA of a target polypeptide or they can interact with the polypeptide itself. Often functional nucleic acids are designed to interact with other nucleic acids based on sequence homology between the target molecule and the functional nucleic acid molecule. In other situations, the specific recognition between the functional nucleic acid molecule and the target molecule is not based on sequence homology between the functional nucleic acid molecule and the target molecule, but rather is based on the formation of tertiary structure that allows specific recognition to take place.

Therefore the compositions can include one or more functional nucleic acids designed to reduce expression of a gene, or a gene product thereof. For example, the functional nucleic acid or polypeptide can be designed to target and reduce or inhibit expression or translation of an mRNA; or to reduce or inhibit expression, reduce activity, or increase degradation of a protein. In some embodiments, the composition includes a vector suitable for in vivo expression of the functional nucleic acid.

i. Antisense

The functional nucleic acids can be or encode antisense molecules. Antisense molecules are designed to interact with a target nucleic acid molecule through either canonical or non-canonical base pairing. The interaction of the antisense molecule and the target molecule is designed to promote the destruction of the target molecule through, for example, RNAse H mediated RNA-DNA hybrid degradation. Alternatively the antisense molecule is designed to interrupt a processing function that normally would take place on the target molecule, such as transcription or replication. Antisense molecules can be designed based on the sequence of the target molecule. There are numerous methods for optimization of antisense efficiency by finding the most accessible regions of the target molecule. Exemplary methods include *in vitro* selection experiments and DNA modification studies using DMS and DEPC. It is preferred that antisense

CA 03149421 2022-2-24

molecules bind the target molecule with a dissociation constant (K_d) less than or equal to 10^{-6}, 10^{-8}, 10^{-10}, or 10^{-12}.

ii. RNA Interference

In some embodiments, the functional nucleic acids induce gene
5 silencing through RNA interference. Gene expression can also be effectively silenced in a highly specific manner through RNA interference (RNAi). This silencing was originally observed with the addition of double stranded RNA (dsRNA) (Fire, *et al.* (1998) Nature, 391:806-11; Napoli, *et al.* (1990) Plant Cell 2:279-89; Hannon, (2002) Nature, 418:244-51). Once dsRNA enters a
10 cell, it is cleaved by an RNase III –like enzyme, Dicer, into double stranded small interfering RNAs (siRNA) 21-23 nucleotides in length that contains 2 nucleotide overhangs on the 3' ends (Elbashir, *et al.* (2001) Genes Dev., 15:188-200; Bernstein, *et al.* (2001) Nature, 409:363-6; Hammond, *et al.* (2000) Nature, 404:293-6). In an ATP dependent step, the siRNAs become
15 integrated into a multi-subunit protein complex, commonly known as the RNAi induced silencing complex (RISC), which guides the siRNAs to the target RNA sequence (Nykanen, *et al.* (2001) Cell, 107:309-21). At some point the siRNA duplex unwinds, and it appears that the antisense strand remains bound to RISC and directs degradation of the complementary
20 mRNA sequence by a combination of endo and exonucleases (Martinez, *et al.* (2002) Cell, 110:563-74). However, the effect of iRNA or siRNA or their use is not limited to any type of mechanism.

Short Interfering RNA (siRNA) is a double-stranded RNA that can induce sequence-specific post-transcriptional gene silencing, thereby
25 decreasing or even inhibiting gene expression. In one example, a siRNA triggers the specific degradation of homologous RNA molecules, such as mRNAs, within the region of sequence identity between both the siRNA and the target RNA. For example, WO 02/44321 discloses siRNAs capable of sequence-specific degradation of target mRNAs when base-paired with 3'
30 overhanging ends, herein incorporated by reference for the method of making these siRNAs.

Sequence specific gene silencing can be achieved in mammalian cells using synthetic, short double-stranded RNAs that mimic the siRNAs produced by the enzyme dicer (Elbashir, *et al.* (2001) Nature, 411:494 498)

CA 03149421 2022-2-24

(Ui-Tei, *et al.* (2000) FEBS Lett 479:79-82). siRNA can be chemically or *in vitro*-synthesized or can be the result of short double-stranded hairpin-like RNAs (shRNAs) that are processed into siRNAs inside the cell. Synthetic siRNAs are generally designed using algorithms and a conventional
5 DNA/RNA synthesizer. Suppliers include Ambion (Austin, Texas), ChemGenes (Ashland, Massachusetts), Dharmacon (Lafayette, Colorado), Glen Research (Sterling, Virginia), MWB Biotech (Esbersberg, Germany), Proligo (Boulder, Colorado), and Qiagen (Vento, The Netherlands). siRNA can also be synthesized *in vitro* using kits such as Ambion's SILENCER®
10 siRNA Construction Kit.

The production of siRNA from a vector is more commonly done through the transcription of a short hairpin RNAse (shRNAs). Kits for the production of vectors having shRNA are available, such as, for example, Imgenex's GENESUPPRESSOR™ Construction Kits and Invitrogen's
15 BLOCK-IT™ inducible RNAi plasmid and lentivirus vectors.

In some embodiment, the functional nucleic acid is siRNA, shRNA, miRNA. In some embodiments, the composition includes a vector expressing the functional nucleic acid.

iii. Aptamers

20 The functional nucleic acids can be or encode an aptamer. Aptamers are molecules that interact with a target molecule, preferably in a specific way. Typically aptamers are small nucleic acids ranging from 15-50 bases in length that fold into defined secondary and tertiary structures, such as stem-loops or G-quartets. Aptamers can bind small molecules, such as ATP and
25 theophiline, as well as large molecules, such as reverse transcriptase and thrombin. Aptamers can bind very tightly with K_d's from the target molecule of less than 10^{-12} M. It is preferred that the aptamers bind the target molecule with a K_d less than 10^{-6},
10^{-8}, 10^{-10}, or 10^{-12}. Aptamers can bind the target molecule with a very high
30 degree of specificity. For example, aptamers have been isolated that have greater than a 10,000 fold difference in binding affinities between the target molecule and another molecule that differ at only a single position on the molecule. It is preferred that the aptamer have a K_d with the target molecule at least 10, 100, 1000, 10,000, or 100,000 fold lower than the K_d with a

CA 03149421 2022-2-24

background binding molecule. It is preferred when doing the comparison for a molecule such as a polypeptide, that the background molecule be a different polypeptide.

iv. Ribozymes

5 The functional nucleic acids can be or encode ribozymes. Ribozymes are nucleic acid molecules that are capable of catalyzing a chemical reaction, either intramolecularly or intermolecularly. It is preferred that the ribozymes catalyze intermolecular reactions. There are a number of different types of ribozymes that catalyze nuclease or nucleic acid polymerase type reactions 10 which are based on ribozymes found in natural systems, such as hammerhead ribozymes. There are also a number of ribozymes that are not found in natural systems, but which have been engineered to catalyze specific reactions *de novo*. Preferred ribozymes cleave RNA or DNA substrates, and more preferably cleave RNA substrates. Ribozymes typically cleave nucleic 15 acid substrates through recognition and binding of the target substrate with subsequent cleavage. This recognition is often based mostly on canonical or non-canonical base pair interactions. This property makes ribozymes particularly good candidates for target specific cleavage of nucleic acids because recognition of the target substrate is based on the target substrates 20 sequence.

v. External Guide Sequences

The functional nucleic acids can be or encode external guide sequences. External guide sequences (EGSs) are molecules that bind a target nucleic acid molecule forming a complex, which is recognized by RNase P, 25 which then cleaves the target molecule. EGSs can be designed to specifically target a RNA molecule of choice. RNAse P aids in processing transfer RNA (tRNA) within a cell. Bacterial RNAse P can be recruited to cleave virtually any RNA sequence by using an EGS that causes the target RNA:EGS complex to mimic the natural tRNA substrate. Similarly, 30 eukaryotic EGS/RNAse P-directed cleavage of RNA can be utilized to cleave desired targets within eukaryotic cells. Representative examples of how to make and use EGS molecules to facilitate cleavage of a variety of different target molecules are known in the art.

CA 03149421 2022-2-24

Methods of making and using vectors for *in vivo* expression of functional nucleic acids such as antisense oligonucleotides, siRNA, shRNA, miRNA, EGSs, ribozymes, and aptamers are known in the art.

vi. Cyclic Dinucleotides

5 The functional nucleic acids can be or encode a cyclic dinucleotide. Cyclic dinucleotides bind directly to the STING adaptor protein, resulting in production of IFN-β (Zhang, et al., Mol Cell., 51(2):226-35 (2013). doi: 10.1016/j.molcel.2013.05.022.). Several canonical and noncanonical dinucleotides are known in the art, and include, but are not limited to, 2'3'-
10 cGAMP , 2'3'-cGAMP , 3'3'-cGAMP, c-di-AMP, c-di-GMP, cAIMP (CL592), cAIMP Difluor (CL614), cAIM(PS)2 Difluor (Rp/Sp) (CL656), 2'2'-cGAMP, 2'3'-cGAM(PS)2 (Rp/Sp), 3'3'-cGAMP Fluorinated, c-di-AMP Fluorinated, 2'3'-c-di-AMP, 2'3'-c-di-AM(PS)2 (Rp,Rp), 2'3'-c-di-AM(PS)2 (Rp,Rp), c-di-GMP Fluorinated, 2'3'-c-di-GMP, c-di-IMP,
15 DMXAA.

vii. Immunostimulatory Oligonucleotides

In some embodiments, the functional nucleic acids can be or encode an oligonucleotide ligand. Examples include, but are not limited to, pattern recognition receptors (PRRs) ligands.

20 Examples of PRRs include the Toll-like family of signaling molecules that play a role in the initiation of innate immune responses and also influence the later and more antigen specific adaptive immune responses. Therefore, the oligonucleotide can serve as a ligand for a Toll-like family signaling molecule, such as Toll-Like Receptor 9 (TLR9).

25 For example, unmethylated CpG sites can be detected by TLR9 on plasmacytoid dendritic cells and B cells in humans (Zaida, et al., *Infection and Immunity*, 76(5):2123-2129, (2008)). Therefore, the sequence of oligonucleotide can include one or more unmethylated cytosine-guanine (CG or CpG, used interchangeably) dinucleotide motifs. The 'p' refers to the
30 phosphodiester backbone of DNA, however, in some embodiments, oligonucleotides including CG can have a modified backbone, for example a phosphorothioate (PS) backbone.

In some embodiments, an oligonucleotide can contain more than one CG dinucleotide, arranged either contiguously or separated by intervening nucleotide(s). The CpG motif(s) can be in the interior of the oligonucleotide sequence. Numerous nucleotide sequences stimulate TLR9 with variations
5 in the number and location of CG dinucleotide(s), as well as the precise base sequences flanking the CG dimers.

Typically, CG ODNs are classified based on their sequence, secondary structures, and effect on human peripheral blood mononuclear cells (PBMCs). The five classes are Class A (Type D), Class B (Type K),
10 Class C, Class P, and Class S (Vollmer, J & Krieg, AM, *Advanced drug delivery reviews* 61(3): 195–204 (2009), incorporated herein by reference). CG ODNs can stimulate the production of Type I interferons (e.g., IFNα) and induce the maturation of dendritic cells (DCs). Some classes of ODNs are also strong activators of natural killer (NK) cells through indirect
15 cytokine signaling. Some classes are strong stimulators of human B cell and monocyte maturation (Weiner, GL, PNAS USA 94(20): 10833-7 (1997); Dalpke, AH, *Immunology* 106(1): 102-12 (2002); Hartmann, G, *J of Immun.* 164(3):1617-2 (2000), each of which is incorporated herein by reference).

Other PRR Toll-like receptors include TLR3, and TLR7 which may
20 recognize double-stranded RNA, single-stranded and short double-stranded RNAs, respectively, and retinoic acid-inducible gene I (RIG-I)-like receptors, namely RIG-I and melanoma differentiation-associated gene 5 (MDA5), which are best known as RNA-sensing receptors in the cytosol.

RIG-I (retinoic-acid-inducible protein 1, also known as Ddx58) and
25 MDA-5 (melanoma-differentiation-associated gene 5, also known as Ifih1 or Helicard) are cytoplasmic RNA helicases that belong to the RIG-I-like receptors (RLRs) family and are critical for host antiviral responses.

RIG-I and MDA-5 sense double-stranded RNA (dsRNA), a replication intermediate for RNA viruses, and signal through the
30 mitochondrial antiviral signaling protein MAVS (also known as IPS-1, VISA or Cardif), leading to production of type-I interferons (IFN-α and IFN-β).

RIG-I detects viral RNA that exhibit an uncapped 5'-di/triphosphate end and a short blunt-ended double stranded potion, two essential features facilitating discrimination from self-RNAs. The features of MDA-5

CA 03149421 2022-2-24

physiological ligands have not been fully characterized yet. However, it is admitted that RIG-I and MDA-5 exhibit a different dependency for the length of dsRNAs: RIG-I selectively binds short dsRNA while MDA-5 selectively binds long dsRNA. Consistent with this, RIG-I and MDA-5 bind
5 Poly(I:C), a synthetic dsRNA analog, with different length predilection.

Under some circumstances, RIG-I can also sense dsDNA indirectly. Viral dsDNA can be transcribed by the RNA polymerase III into dsRNA with a 5'-triphosphate moiety. Poly(dA:dT), a synthetic analog of B-form DNA, thus constitutes another RIG-I ligand.

10 Exemplary RIG-I ligands include, but are not limited to, 5'ppp-dsRNA, a specific agonist of RIG-I; 3p-hpRNA, a specific agonist of RIG-I; Poly(I:C)/LyoVec complexes that are recognized by RIG-I and/or MDA-5 depending of the size of poly(I:C); Poly(dA:dT)/LyoVec complexes that are indirectly recognized by RIG-I.

15 In some embodiments, the oligonucleotide contains a functional ligand for TLR3, TLR7, TLR8, TLR9, or RIG-I-like receptors, or combinations thereof.

Examples of immunostimulatory oligonucleotides, and methods of making them are known in the art and commercially available, see for
20 example, Bodera, P. *Recent Pat Inflamm Allergy Drug Discov.* 5(1):87-93 (2011), incorporated herein by reference.

3. Composition of the Cargo

The disclosed nucleic acid cargo can be or include DNA or RNA nucleotides which typically include a heterocyclic base (nucleic acid base), a
25 sugar moiety attached to the heterocyclic base, and a phosphate moiety which esterifies a hydroxyl function of the sugar moiety. The principal naturally-occurring nucleotides include uracil, thymine, cytosine, adenine and guanine as the heterocyclic bases, and ribose or deoxyribose sugar linked by phosphodiester bonds.

30 In some embodiments, the cargo includes or is composed of nucleotide analogs that have been chemically modified to improve stability, half-life, or specificity or affinity for a target receptor, relative to a DNA or RNA counterpart. The chemical modifications include chemical modification of nucleobases, sugar moieties, nucleotide linkages, or

CA 03149421 2022-2-24

combinations thereof. As used herein 'modified nucleotide" or "chemically modified nucleotide" defines a nucleotide that has a chemical modification of one or more of the heterocyclic base, sugar moiety or phosphate moiety constituents. In some embodiments, the charge of the modified nucleotide is reduced compared to DNA or RNA of the same nucleobase sequence. For example, the oligonucleotide can have low negative charge, no charge, or positive charge.

Typically, nucleoside analogs support bases capable of hydrogen bonding by Watson-Crick base pairing to standard polynucleotide bases, where the analog backbone presents the bases in a manner to permit such hydrogen bonding in a sequence-specific fashion between the oligonucleotide analog molecule and bases in a standard polynucleotide (e.g., single-stranded RNA or single-stranded DNA). In some embodiments, the analogs have a substantially uncharged, phosphorus containing backbone.

a. Heterocyclic Bases

The principal naturally-occurring nucleotides include uracil, thymine, cytosine, adenine and guanine as the heterocyclic bases. The cargo can include chemical modifications to their nucleobase constituents. Chemical modifications of heterocyclic bases or heterocyclic base analogs may be effective to increase the binding affinity or stability in binding a target sequence. Chemically-modified heterocyclic bases include, but are not limited to, inosine, 5-(1-propynyl) uracil (pU), 5-(1-propynyl) cytosine (pC), 5-methylcytosine, 8-oxo-adenine, pseudocytosine, pseudoisocytosine, 5 and 2-amino-5-(2'-deoxy-.beta.-D-ribofuranosyl)pyridine (2-aminopyridine), and various pyrrolo- and pyrazolopyrimidine derivatives.

b. Sugar Modifications

Cargo can also contain nucleotides with modified sugar moieties or sugar moiety analogs. Sugar moiety modifications include, but are not limited to, 2'-O-aminoetoxy, 2'-O-amonioethyl (2'-OAE), 2'-O-methoxy, 2'-O-methyl, 2-guanidoethyl (2'-OGE), 2'-O,4'-C-methylene (LNA), 2'-O-(methoxyethyl) (2'-OME) and 2'-O-(N-(methyl)acetamido) (2'-OMA). 2'-O-aminoethyl sugar moiety substitutions are especially preferred because they are protonated at neutral pH and thus suppress the charge repulsion between the TFO and the target duplex. This modification stabilizes the C3'-endo

CA 03149421 2022-2-24

conformation of the ribose or dexyribose and also forms a bridge with the i-1 phosphate in the purine strand of the duplex.

In some embodiments, the nucleic acid is a morpholino oligonucleotide. Morpholino oligonucleotides are typically composed of two more morpholino monomers containing purine or pyrimidine base-pairing moieties effective to bind, by base-specific hydrogen bonding, to a base in a polynucleotide, which are linked together by phosphorus-containing linkages, one to three atoms long, joining the morpholino nitrogen of one monomer to the 5' exocyclic carbon of an adjacent monomer. The purine or pyrimidine base-pairing moiety is typically adenine, cytosine, guanine, uracil or thymine. The synthesis, structures, and binding characteristics of morpholino oligomers are detailed in U.S. Patent Nos. 5,698,685, 5,217,866, 5,142,047, 5,034,506, 5,166,315, 5,521,063, and 5,506,337.

Important properties of the morpholino-based subunits typically include: the ability to be linked in a oligomeric form by stable, uncharged backbone linkages; the ability to support a nucleotide base (e.g. adenine, cytosine, guanine, thymidine, uracil or inosine) such that the polymer formed can hybridize with a complementary-base target nucleic acid, including target RNA, with high T_m, even with oligomers as short as 10-14 bases; the ability of the oligomer to be actively transported into mammalian cells; and the ability of an oligomer:RNA heteroduplex to resist RNAse degradation.

In some embodiments, oligonucleotides employ morpholino-based subunits bearing base-pairing moieties, joined by uncharged linkages, as described above.

c. Internucleotide Linkages

Oligonucleotides are connected by an internucleotide bond that refers to a chemical linkage between two nucleoside moieties. Modifications to the phosphate backbone of DNA or RNA oligonucleotides may increase the binding affinity or stability oligonucleotides, or reduce the susceptibility of oligonucleotides nuclease digestion. Cationic modifications, including, but not limited to, diethyl-ethylenediamide (DEED) or dimethyl-aminopropylamine (DMAP) may be especially useful due to decrease electrostatic repulsion between the oligonucleotide and a target. Modifications of the phosphate backbone may also include the substitution

CA 03149421 2022-2-24

of a sulfur atom for one of the non-bridging oxygens in the phosphodiester linkage. This substitution creates a phosphorothioate internucleoside linkage in place of the phosphodiester linkage. Oligonucleotides containing phosphorothioate internucleoside linkages have been shown to be more
5 stable *in vivo*.

Examples of modified nucleotides with reduced charge include modified internucleotide linkages such as phosphate analogs having achiral and uncharged intersubunit linkages (e.g., Sterchak, E. P. et al., *Organic. Chem.*, 52:4202, (1987)), and uncharged morpholino-based polymers having
10 achiral intersubunit linkages (see, e.g., U.S. Pat. No. 5,034,506), as discussed above. Some internucleotide linkage analogs include morpholidate, acetal, and polyamide-linked heterocycles.

In another embodiment, the cargo are composed of locked nucleic acids. Locked nucleic acids (LNA) are modified RNA nucleotides (see, for
15 example, Braasch, et al., *Chem. Biol.*, 8(1):1-7 (2001)). LNAs form hybrids with DNA which are more stable than DNA/DNA hybrids, a property similar to that of peptide nucleic acid (PNA)/DNA hybrids. Therefore, LNA can be used just as PNA molecules would be. LNA binding efficiency can be increased in some embodiments by adding positive charges to it.
20 Commercial nucleic acid synthesizers and standard phosphoramidite chemistry are used to make LNAs.

In some embodiments, the cargo are composed of peptide nucleic acids. Peptide nucleic acids (PNAs) are synthetic DNA mimics in which the phosphate backbone of the oligonucleotide is replaced in its entirety by
25 repeating N-(2-aminoethyl)-glycine units and phosphodiester bonds are typically replaced by peptide bonds. The various heterocyclic bases are linked to the backbone by methylene carbonyl bonds. PNAs maintain spacing of heterocyclic bases that is similar to conventional DNA oligonucleotides, but are achiral and neutrally charged molecules. Peptide
30 nucleic acids are composed of peptide nucleic acid monomers.

Other backbone modifications include peptide and amino acid variations and modifications. Thus, the backbone constituents of oligonucleotides such as PNA may be peptide linkages, or alternatively, they may be non-peptide peptide linkages. Examples include acetyl caps, amino

CA 03149421 2022-2-24

spacers such as 8-amino-3,6-dioxaoctanoic acid (referred to herein as O-linkers), amino acids such as lysine are particularly useful if positive charges are desired in the PNA, and the like. Methods for the chemical assembly of PNAs are well known. See, for example, U.S. Patent Nos. 5,539,082,
5 5,527,675, 5,623,049, 5,714,331, 5,736,336, 5,773,571 and 5,786,571.

Cargo optionally include one or more terminal residues or modifications at either or both termini to increase stability, and/or affinity of the oligonucleotide for its target. Commonly used positively charged moieties include the amino acids lysine and arginine, although other
10 positively charged moieties may also be useful. Cargo may further be modified to be end capped to prevent degradation using a propylamine group. Procedures for 3' or 5' capping oligonucleotides are well known in the art.

In some embodiments, the nucleic acid can be single stranded or
15 double stranded.

C. Pharmaceutical Compositions

The compositions can be used therapeutically in combination with a pharmaceutically acceptable carrier.

The compositions including nucleic acid cargo complexed with 3E10
20 antibody are preferably employed for therapeutic uses in combination with a suitable pharmaceutical carrier. Such compositions include an effective amount of the composition, and a pharmaceutically acceptable carrier or excipient.

The compositions may be in a formulation for administration
25 topically, locally or systemically in a suitable pharmaceutical carrier. Remington's Pharmaceutical Sciences, 15th Edition by E. W. Martin (Mark Publishing Company, 1975), discloses typical carriers and methods of preparation. The complexes may also be encapsulated in suitable biocompatible particles formed of biodegradable or non-biodegradable
30 polymers or proteins or liposomes for targeting to cells. Such systems are well known to those skilled in the art. In some embodiments, the complexes are encapsulated in nanoparticles.

Formulations for injection may be presented in unit dosage form, e.g., in ampules or in multi-dose containers, optionally with an added

CA 03149421 2022-2-24

preservative. The compositions may take such forms as sterile aqueous or nonaqueous solutions, suspensions and emulsions, which can be isotonic with the blood of the subject in certain embodiments. Examples of nonaqueous solvents are polypropylene glycol, polyethylene glycol, vegetable oil such as olive oil, sesame oil, coconut oil, arachis oil, peanut oil, mineral oil, injectable organic esters such as ethyl oleate, or fixed oils including synthetic mono or di-glycerides. Aqueous carriers include water, alcoholic/aqueous solutions, emulsions or suspensions, including saline and buffered media. Parenteral vehicles include sodium chloride solution, 1,3-butandiol, Ringer's dextrose, dextrose and sodium chloride, lactated Ringer's or fixed oils. Intravenous vehicles include fluid and nutrient replenishers, and electrolyte replenishers (such as those based on Ringer's dextrose).The materials may be in solution, emulsions, or suspension (for example, incorporated into particles, liposomes, or cells). Typically, an appropriate amount of a pharmaceutically-acceptable salt is used in the formulation to render the formulation isotonic. Trehalose, typically in the amount of 1-5%, may be added to the pharmaceutical compositions. The pH of the solution is preferably from about 5 to about 8, and more preferably from about 7 to about 7.5.

Pharmaceutical compositions may include carriers, thickeners, diluents, buffers, preservatives, and surface-active agents. Carrier formulation can be found in Remington's Pharmaceutical Sciences, Mack Publishing Co., Easton, Pa. Those of skill in the art can readily determine the various parameters for preparing and formulating the compositions without resort to undue experimentation.

The compositions alone or in combination with other suitable components, can also be made into aerosol formulations (i.e., they can be "nebulized") to be administered via inhalation. Aerosol formulations can be placed into pressurized acceptable propellants, such as dichlorodifluoromethane, propane, nitrogen, and air. For administration by inhalation, the compounds are delivered in the form of an aerosol spray presentation from pressurized packs or a nebulizer, with the use of a suitable propellant.

CA 03149421 2022-2-24

In some embodiments, the include pharmaceutically acceptable carriers with formulation ingredients such as salts, carriers, buffering agents, emulsifiers, diluents, excipients, chelating agents, preservatives, solubilizers, or stabilizers.

5 The nucleic acids may be conjugated to lipophilic groups like cholesterol and lauric and lithocholic acid derivatives with C32 functionality to improve cellular uptake. For example, cholesterol has been demonstrated to enhance uptake and serum stability of siRNA *in vitro* (Lorenz, et al., *Bioorg. Med. Chem. Lett.,* 14(19):4975-4977 (2004)) and *in vivo* (Soutschek, 10 et al., *Nature,* 432(7014):173-178 (2004)). In addition, it has been shown that binding of steroid conjugated oligonucleotides to different lipoproteins in the bloodstream, such as LDL, protect integrity and facilitate biodistribution (Rump, et al., *Biochem. Pharmacol.,* 59(11):1407-1416 (2000)). Other groups that can be attached or conjugated to the nucleic acids 15 described above to increase cellular uptake, include acridine derivatives; cross-linkers such as psoralen derivatives, azidophenacyl, proflavin, and azidoproflavin; artificial endonucleases; metal complexes such as EDTA-Fe(II) and porphyrin-Fe(II); alkylating moieties; nucleases such as alkaline phosphatase; terminal transferases; abzymes; cholesteryl moieties; lipophilic 20 carriers; peptide conjugates; long chain alcohols; phosphate esters; radioactive markers; non-radioactive markers; carbohydrates; and polylysine or other polyamines. U.S. Patent No. 6,919,208 to Levy, et al., also describes methods for enhanced delivery. These pharmaceutical formulations may be manufactured in a manner that is itself known, e.g., by 25 means of conventional mixing, dissolving, granulating, levigating, emulsifying, encapsulating, entrapping or lyophilizing processes.

Further carriers include sustained release preparations such as semi-permeable matrices of solid hydrophobic polymers containing the complexes, which matrices are in the form of shaped particles, e.g., films, 30 liposomes or microparticles. Implantation includes inserting implantable drug delivery systems, e.g., microspheres, hydrogels, polymeric reservoirs, cholesterol matrixes, polymeric systems, e.g., matrix erosion and/or diffusion systems and non-polymeric systems. Inhalation includes administering the composition with an aerosol in an inhaler, either alone or attached to a carrier

CA 03149421 2022-2-24

that can be absorbed. For systemic administration, it may be preferred that the composition is encapsulated in liposomes.

The compositions may be delivered in a manner which enables tissue-specific uptake of the agent and/or nucleotide delivery system, using invasive devices such as vascular or urinary catheters, and using interventional devices such as stents having drug delivery capability and configured as expansive devices or stent grafts.

The formulations may be delivered using a bioerodible implant by way of diffusion or by degradation of the polymeric matrix. In certain embodiments, the administration of the formulation may be designed to result in sequential exposures to the composition, over a certain time period, for example, hours, days, weeks, months or years. This may be accomplished, for example, by repeated administrations of a formulation or by a sustained or controlled release delivery system in which the compositions are delivered over a prolonged period without repeated administrations.

Other delivery systems suitable include time-release, delayed release, sustained release, or controlled release delivery systems. Such systems may avoid repeated administrations in many cases, increasing convenience to the subject and the physician. Many types of release delivery systems are available and known to those of ordinary skill in the art. They include, for example, polymer-based systems such as polylactic and/or polyglycolic acids, polyanhydrides, polycaprolactones, copolyoxalates, polyesteramides, polyorthoesters, polyhydroxybutyric acid, and/or combinations of these. Microcapsules of the foregoing polymers containing nucleic acids are described in, for example, U.S. Patent No. 5,075,109. Other examples include non-polymer systems that are lipid-based including sterols such as cholesterol, cholesterol esters, and fatty acids or neutral fats such as mono-, di- and triglycerides; hydrogel release systems; liposome-based systems; phospholipid based-systems; silastic systems; peptide based systems; wax coatings; compressed tablets using conventional binders and excipients; or partially fused implants. The formulation may be as, for example, microspheres, hydrogels, polymeric reservoirs, cholesterol matrices, or polymeric systems. In some embodiments, the system may allow sustained

CA 03149421 2022-2-24

or controlled release of the composition to occur, for example, through control of the diffusion or erosion/degradation rate of the formulations containing the complexes.

Complexes include nucleic acid cargo and antibody, and compositions thereof can be formulated for pulmonary or mucosal administration. The administration can include delivery of the composition to the lungs, nasal, oral (sublingual, buccal), vaginal, or rectal mucosa. The term aerosol as used herein refers to any preparation of a fine mist of particles, which can be in solution or a suspension, whether or not it is produced using a propellant. Aerosols can be produced using standard techniques, such as ultrasonication or high-pressure treatment.

For administration via the upper respiratory tract, the formulation can be formulated into a solution, e.g., water or isotonic saline, buffered or un-buffered, or as a suspension, for intranasal administration as drops or as a spray. Preferably, such solutions or suspensions are isotonic relative to nasal secretions and of about the same pH, ranging e.g., from about pH 4.0 to about pH 7.4 or, from pH 6.0 to pH 7.0. Buffers should be physiologically compatible and include, simply by way of example, phosphate buffers.

The complexes can be delivered to the target cells using a particle delivery vehicle. Nanoparticles generally refers to particles in the range of between 500 nm to less than 0.5 nm, preferably having a diameter that is between 50 and 500 nm, more preferably having a diameter that is between 50 and 300 nm. Cellular internalization of polymeric particles is highly dependent upon their size, with nanoparticulate polymeric particles being internalized by cells with much higher efficiency than micoparticulate polymeric particles. For example, Desai, et al. have demonstrated that about 2.5 times more nanoparticles that are 100 nm in diameter are taken up by cultured Caco-2 cells as compared to microparticles having a diameter on 1 μM (Desai, et al., *Pharm. Res.*, 14:1568-73 (1997)). Nanoparticles also have a greater ability to diffuse deeper into tissues *in vivo*.

In some embodiments, the delivery vehicle is a dendrimer.

Examples of preferred biodegradable polymers include synthetic polymers that degrade by hydrolysis such as poly(hydroxy acids), such as

CA 03149421 2022-2-24

polymers and copolymers of lactic acid and glycolic acid, other degradable polyesters, polyanhydrides, poly(ortho)esters, polyesters, polyurethanes, poly(butic acid), poly(valeric acid), poly(caprolactone), poly(hydroxyalkanoates), poly(lactide-co-caprolactone), and poly(amine-co-ester) polymers, such as those described in Zhou, et al., *Nature Materials*, 11:82-90 (2012) and WO 2013/082529, U.S. Published Application No. 2014/0342003, and PCT/US2015/061375.

In some embodiments, particularly those for targeting T cell in vivo, for example, for in vivo production of CAR T cells, immune cell or T cell markers such as CD3, CD7, or CD8, or markers of a target tissue such as the liver, can be targeted. For example, anti-CD8 antibodies and anti-CD3 Fab fragments have both been used to target T cells in vivo (Pfeiffer, et al., *EMBO Mol Med.*, 10(11) (2018). pii: e9158. doi: 10.15252/emmm.201809158., Smith, et al., *Nat Nanotechnol.*, 12(8):813-820 (2017). doi: 10.1038/nnano.2017.57). Thus, in some embodiments, the particle or other delivery vehicle includes a targeting moiety specific for CD3, CD7, CD8, or another immune cell (e.g., T cell) marker, or a marker for a specific tissue such as the thymus, spleen, or liver. The binding moiety can be, for example, an antibody or antigen binding fragment thereof.

Targeting moieties can be associated with, linked, conjugated, or otherwise attached directly or indirectly to a nanoparticle or other delivery vehicle thereof. Targeting molecules can be proteins, peptides, nucleic acid molecules, saccharides or polysaccharides that bind to a receptor or other molecule on the surface of a targeted cell. The degree of specificity and the avidity of binding to the graft can be modulated through the selection of the targeting molecule.

Examples of moieties include, for example, targeting moieties which provide for the delivery of molecules to specific cells, e.g., antibodies to hematopoietic stem cells, $CD34^{+}$ cells, T cells or any other preferred cell type, as well as receptor and ligands expressed on the preferred cell type. Preferably, the moieties target hematopoeitic stem cells. Examples of molecules targeting extracellular matrix ("ECM") include glycosaminoglycan ("GAG") and collagen. In one embodiment, the external surface of polymer particles may be modified to enhance the ability of the

CA 03149421 2022-2-24

particles to interact with selected cells or tissue. The method described above wherein an adaptor element conjugated to a targeting molecule is inserted into the particle is preferred. However, in another embodiment, the outer surface of a polymer micro- or nanoparticle having a carboxy terminus may be linked to targeting molecules that have a free amine terminus.

Other useful ligands attached to polymeric micro- and nanoparticles include pathogen-associated molecular patterns (PAMPs). PAMPs target Toll-like Receptors (TLRs) on the surface of the cells or tissue, or signal the cells or tissue internally, thereby potentially increasing uptake. PAMPs conjugated to the particle surface or co-encapsulated may include: unmethylated CpG DNA (bacterial), double-stranded RNA (viral), lipopolysacharride (bacterial), peptidoglycan (bacterial), lipoarabinomannin (bacterial), zymosan (yeast), mycoplasmal lipoproteins such as MALP-2 (bacterial), flagellin (bacterial) poly(inosinic-cytidylic) acid (bacterial), lipoteichoic acid (bacterial) or imidazoquinolines (synthetic).

In another embodiment, the outer surface of the particle may be treated using a mannose amine, thereby mannosylating the outer surface of the particle. This treatment may cause the particle to bind to the target cell or tissue at a mannose receptor on the antigen presenting cell surface. Alternatively, surface conjugation with an immunoglobulin molecule containing an Fc portion (targeting Fc receptor), heat shock protein moiety (HSP receptor), phosphatidylserine (scavenger receptors), and lipopolysaccharide (LPS) are additional receptor targets on cells or tissue.

Lectins that can be covalently attached to micro- and nanoparticles to render them target specific to the mucin and mucosal cell layer.

The choice of targeting moiety will depend on the method of administration of the nanoparticle composition and the cells or tissues to be targeted. The targeting molecule may generally increase the binding affinity of the particles for cell or tissues or may target the nanoparticle to a particular tissue in an organ or a particular cell type in a tissue. . In some embodiments, the targeting moiety targets the thymus, spleen, or cancer cells

The covalent attachment of any of the natural components of mucin in either pure or partially purified form to the particles would decrease the surface tension of the bead-gut interface and increase the solubility of the

CA 03149421 2022-2-24

bead in the mucin layer. The attachment of polyamino acids containing extra pendant carboxylic acid side groups, e.g., polyaspartic acid and polyglutamic acid, increases bioadhesiveness. Using polyamino acids in the 15,000 to 50,000 kDa molecular weight range yields chains of 120 to 425 amino acid residues attached to the surface of the particles. The polyamino chains increase bioadhesion by means of chain entanglement in mucin strands as well as by increased carboxylic charge.

III. Methods of Use

Methods for using 3E10 to enhance delivery of nucleic acid constructs are provided. Typically an effective amount of 3E10 antibody is first contacted with a nucleic acid cargo whose delivery into cells is desired. For example, the nucleic acid cargo and antibody can be mixed in solution for sufficient time for the nucleic acid cargo and antibody to form complexes. Next, the mixture is contacted with cells. In other embodiments, the cargo and antibody are added to a solution containing or otherwise bathing cells, and the complexes are formed in the presence of the cells. The complexes can be contacted with cells *in vitro, ex vivo,* or *in vivo*. Thus, in some embodiments, the solution of complexes is added to the cells in culture or injected into an animal to be treated.

It is believed that the antibody helps deliver the nucleic acid into cell nuclei, and then alters the function of the RAD51 pathway which can promote gene editing by the donor DNA. The approach has no sequence limitations to the design of the nucleic acid cargo. The treatment can be, for example, administration of a mixture of an antibody and nucleic acid cargo to a subject in need thereof by simple IV administration

The compositions and methods can include 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more different nucleic acid constructs formed of RNA, DNA, PNA or other modified nucleic acids, or a combination thereof.

The effective amount or therapeutically effective amount of the composition can be a dosage sufficient to treat, inhibit, or alleviate one or more symptoms of a disease or disorder, or to otherwise provide a desired pharmacologic and/or physiologic effect, for example, reducing, inhibiting, or reversing one or more of the pathophysiological mechanisms underlying a disease or disorder.

CA 03149421 2022-2-24

An effective amount may also be an amount effective to increase the rate, quantity, and/or quality of delivery of the nucleic acid cargo relative to administration of the cargo in the absence of the antibody. The formulation of the composition is made to suit the mode of administration.
5 Pharmaceutically acceptable carriers are determined in part by the particular composition being administered, as well as by the particular method used to administer the composition. Accordingly, there is a wide variety of suitable formulations of pharmaceutical compositions containing the complexes. The precise dosage will vary according to a variety of factors such as subject-
10 dependent variables (e.g., age, immune system health, clinical symptoms etc.).

The composition can be administered or otherwise contacted with target cells once, twice, or three times daily; one, two, three, four, five, six, seven times a week, one, two, three, four, five, six, seven or eight times a
15 month. For example, in some embodiments, the composition is administered every two or three days, or on average about 2 to about 4 times about week. Thus, in some embodiments, the composition is administered as part of dosage regimen including two or more separate treatments.

Dosage regimens include maintenance regimens, where the dosage
20 remains the same between two or more administrations, escalation regimens where the dosage increases between two or more administrations, de-escalation regimens, where the dosage decreases between two or more administrations, or a combination thereof.

In some embodiments, the first dose can be a low dose. Dose
25 escalation can be continued until a satisfactory biochemical or clinical response is reached. The clinical response will depend on the disease or disorder being treated, and/or the desired outcome. In some embodiments the dosage may increase until a therapeutic effect is identified, preferably without also inducing undesired toxicity or an acceptably high amount
30 thereof. Next, the dosages can be maintained or steadily reduced to a maintenance dose. The methods can used to standardize, optimize, or customize the dose level, dose frequency, or duration of the therapy.

Generally, prior to administration, particularly for in vivo administration, antibody and nucleic acid are mixed for a period of time at

CA 03149421 2022-2-24

room temperature. In some embodiments, time of complexation ranges from, for example, 1 minute to 30 minutes, or 10 minutes to 20, each inclusive, with a preferred complexation time of about 15 minutes. Antibody dose can range from 0.0001 mg to 1 mg, each inclusive, with a preferred
5 dose of about 0.1 mg. Nucleic acid dose can range from 0.001 µg to 100 µg, inclusive, with a preferred dose of 10 µg. The in vivo data below (e.g., Fig. 6 B) was produced using 0.1 mg 3E10, 10 µg of mRNA, and complexed for 15 minutes.

The Examples below may indicate that DNA cargo may be delivered
10 more generally to multiple tissues and not restricted to tumors, while RNA delivery may be more selective for tumor tissue. Thus, in some embodiments, RNA cargo (e.g., alone) may be selectively delivered to cancer cells or other tumor tissues. In some embodiments, when wider distribution of RNA cargo is desired, the RNA may be mixed with DNA
15 (e.g., carrier DNA) to facilitate delivery to non-cancer/tumor tissues. Carrier DNA can be, for example, plasmid DNA or low molecular weight, from e.g., salmon sperm. In some embodiments, carrier DNA is non-coding DNA. Carrier DNA can be single stranded or double stranded or a combination thereof. In some embodiments, carrier DNA is composed of nucleic acids
20 having 1-10, 1-100, 1-1,000, or 1-10,000 nucleotides in length, or any subrange or integer thereof, or combination thereof. Typically carrier DNA is not conjugated or otherwise covalently attached to the antibody. Typically carrier DNA is co-incubated with cargo nucleic acid (e.g., RNA) and antibody, and co-delivered as a complex therewith. In some embodiments,
25 the carrier DNA is non-coding DNA.

A. *In vitro* and *Ex vivo* Methods

For *in vitro* and *ex vivo* methods, cells are typically contacted with the composition while in culture. For *ex vivo* methods, cells may be isolated from a subject and contacted *ex vivo* with the composition to produce cells
30 containing the cargo nucleic acid(s). In a preferred embodiment, the cells are isolated from the subject to be treated or from a syngenic host. Target cells can be removed from a subject prior to contacting with composition.

CA 03149421 2022-2-24

B. *In vivo* Methods

In some embodiments, *in vivo* delivery of nucleic acid cargo to cells is used for gene editing and/or treatment of a disease or disorder in a subject. The composition, typically including antibody-nucleic acid cargo, can be administered directly to a subject for *in vivo* therapy.

In general, methods of administering compounds, including antibodies, oligonucleotides and related molecules, are well known in the art. In particular, the routes of administration already in use for nucleic acid therapeutics, along with formulations in current use, provide preferred routes of administration and formulation for the donor oligonucleotides described above. Preferably the composition is injected or infused into the animal.

The compositions can be administered by a number of routes including, but not limited to, intravenous, intraperitoneal, intraamniotic, intramuscular, subcutaneous, or topical (sublingual, rectal, intranasal, pulmonary, rectal mucosa, and vaginal), and oral (sublingual, buccal).

In some embodiments, the composition is formulated for pulmonary delivery, such as intranasal administration or oral inhalation. Administration of the formulations may be accomplished by any acceptable method that allows the complexes to reach their targets. The administration may be localized (i.e., to a particular region, physiological system, tissue, organ, or cell type) or systemic, depending on the condition being treated. Compositions and methods for *in vivo* delivery are also discussed in WO 2017/143042.

The methods can also include administering an effective amount of the antibody-nucleic acid complex composition to an embryo or fetus, or the pregnant mother thereof, *in vivo*. In some methods, compositions are delivered *in utero* by injecting and/or infusing the compositions into a vein or artery, such as the vitelline vein or the umbilical vein, or into the amniotic sac of an embryo or fetus. See, e.g., Ricciardi, et al., *Nat Commun*. 2018 Jun 26;9(1):2481. doi: 10.1038/s41467-018-04894-2, and WO 2018/187493.

C. Applications

Nucleic acid cargo, e.g., mRNA, functional nucleic acid, DNA expression constructs, vectors, etc., encoding a polypeptide of interest or functional nucleic acid, can be delivered into cells using a 3E10 antibody, for

CA 03149421 2022-2-24

expression of, or inhibition of, a polypeptide in the cells. The compositions and methods can be used over a range of different applications. Non-limiting examples include CRISPR and gRNA expression vectors +/- editing DNAs, delivery of large DNAs (plasmids and expression vectors), gene replacement and gene therapy, delivery of DNAs and/or RNAs to, for example, generate CAR-T cells in vivo or ex vivo and to simplify CAR-T cell production in vivo or ex vivo, delivery of siRNAs, delivery of mRNAs, etc. Exemplary applications related to gene therapy/gene editing and immunomodulation, particularly chimeric antigen receptor T cell production, are discussed below.

1. Gene Therapy and Editing

In some embodiments, the compositions are used for gene editing.

For example, the methods can be especially useful to treat genetic deficiencies, disorders and diseases caused by mutations in single genes, for example, to correct genetic deficiencies, disorders and diseases caused by point mutations. If the target gene contains a mutation that is the cause of a genetic disorder, then the methods can be used for mutagenic repair that may restore the DNA sequence of the target gene to normal. The target sequence can be within the coding DNA sequence of the gene or within an intron. The target sequence can also be within DNA sequences that regulate expression of the target gene, including promoter or enhancer sequences.

In the methods herein, cells that have been contacted with the complexes may be administered to a subject. The subject may have a disease or disorder such as hemophilia, muscular dystrophy, globinopathies, cystic fibrosis, xeroderma pigmentosum, or lysosomal storage diseases. In such embodiments, gene modification, gene replacement, gene addition, or a combination thereof, may occur in an effective amount to reduce one or more symptoms of the disease or disorder in the subject.

In some embodiments, the DNA cargo includes a nucleic acid encoding a nuclease, a donor oligonucleotide or nucleic acid encoding a donor oligonucleotide, or a combination thereof.

a. Gene Editing Nuclease

Nucleic acid cargos include those that encode an element or elements that induce a single or a double strand break in the target cell's genome, and optionally, but preferable in combination with other elements such as donor

CA 03149421 2022-2-24

oligonucleotides and/or, particularly in the case of CRISPR/Cas, other elements of the system such as gRNA. The compositions can be used, for example, to reduce or otherwise modify expression of a target gene.

i. Strand Break Inducing Elements

5 ***CRISPR/Cas***

In some embodiments, the nucleic acid cargo includes one or more elements of a CRISPR/Cas-mediated genome editing composition, a nucleic acid encoding one or more elements of a CRISPR/Cas-mediated genome editing composition, or a combination thereof. As used herein,
10 CRISPR/Cas-mediated genome editing composition refers to the elements of a CRISPR system needed to carry out CRISPR/Cas-mediated genome editing in a mammalian subject. As discussed in more detail below, CRISPR/Cas-mediated genome editing compositions typically include one or more nucleic acids encoding a crRNA, a tracrRNA (or chimeric thereof also
15 referred to a guide RNA or single guide RNA) and a Cas enzyme, such as Cas9. The CRISPR/Cas-mediated genome editing composition can optionally include a donor polynucleotide that can be recombined into the target cell's genome at or adjacent to the target site (e.g., the site of single or double stand break induced by the Cas9).

20 The CRISPR/Cas system has been adapted for use as gene editing (silencing, enhancing or changing specific genes) for use in eukaryotes (see, for example, Cong, *Science*, 15:339(6121):819–823 (2013) and Jinek, et al., *Science*, 337(6096):816-21 (2012)). By transfecting a cell with the required elements including a cas gene and specifically designed CRISPRs, the
25 organism's genome can be cut and modified at any desired location. Methods of preparing compositions for use in genome editing using the CRISPR/Cas systems are described in detail in WO 2013/176772 and WO 2014/018423, which are specifically incorporated by reference herein in their entireties.

The methods of delivery disclosed herein are suitable for use with
30 numerous variations on the CRISPR/Cas system.

In general, "CRISPR system" refers collectively to transcripts and other elements involved in the expression of or directing the activity of CRISPR-associated ("Cas") genes, including sequences encoding a Cas gene, a tracr (trans-activating CRISPR) sequence (e.g., tracrRNA or an

CA 03149421 2022-2-24

active partial tracrRNA), a tracr-mate sequence (encompassing a "direct repeat" and a tracrRNA-processed partial direct repeat in the context of an endogenous CRISPR system), a guide sequence (also referred to as a "spacer" in the context of an endogenous CRISPR system), or other
5 sequences and transcripts from a CRISPR locus. One or more tracr mate sequences operably linked to a guide sequence (e.g., direct repeat-spacer-direct repeat) can also be referred to as pre-crRNA (pre-CRISPR RNA) before processing or crRNA after processing by a nuclease.

As discussed in more detail below, in some embodiments, a
10 tracrRNA and crRNA are linked and form a chimeric crRNA-tracrRNA hybrid where a mature crRNA is fused to a partial tracrRNA via a synthetic stem loop to mimic the natural crRNA:tracrRNA duplex as described in Cong, *Science*, 15:339(6121):819–823 (2013) and Jinek, et al., *Science*, 337(6096):816-21 (2012)). A single fused crRNA-tracrRNA construct is also
15 referred to herein as a guide RNA or gRNA (or single-guide RNA (sgRNA)). Within an sgRNA, the crRNA portion can be identified as the 'target sequence' and the tracrRNA is often referred to as the 'scaffold'.

In some embodiments, one or more elements of a CRISPR system is derived from a type I, type II, or type III CRISPR system. In some
20 embodiments, one or more elements of a CRISPR system is derived from a particular organism including an endogenous CRISPR system, such as *Streptococcus pyogenes*.

In general, a CRISPR system is characterized by elements that promote the formation of a CRISPR complex at the site of a target sequence
25 (also referred to as a protospacer in the context of an endogenous CRISPR system). In the context of formation of a CRISPR complex, "target sequence" refers to a sequence to which a guide sequence is designed to have complementarity, where hybridization between a target sequence and a guide sequence promotes the formation of a CRISPR complex. A target sequence
30 can be any polynucleotide, such as DNA or RNA polynucleotides. In some embodiments, a target sequence is located in the nucleus or cytoplasm of a cell.

In the target nucleic acid, each protospacer is associated with a protospacer adjacent motif (PAM) whose recognition is specific to individual

CRISPR systems. In the *Streptococcus pyogenes* CRISPR/Cas system, the PAM is the nucleotide sequence NGG. In the *Streptococcus thermophiles* CRISPR/Cas system, the PAM is the nucleotide sequence is NNAGAAW. The tracrRNA duplex directs Cas to the DNA target consisting of the
5 protospacer and the requisite PAM via heteroduplex formation between the spacer region of the crRNA and the protospacer DNA.

Typically, in the context of an endogenous CRISPR system, formation of a CRISPR complex (including a guide sequence hybridized to a target sequence and complexed with one or more Cas proteins) results in
10 cleavage of one or both strands in or near (e.g., within 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50, or more base pairs from) the target sequence. All or a portion of the tracr sequence may also form part of a CRISPR complex, such as by hybridization to all or a portion of a tracr mate sequence that is operably linked to the guide sequence.

15 There are many resources available for helping practitioners determine suitable target sites once a desired DNA target sequence is identified. For example, numerous public resources, including a bioinformatically generated list of about 190,000 potential sgRNAs, targeting more than 40% of human exons, are available to aid practitioners in
20 selecting target sites and designing the associate sgRNA to affect a nick or double strand break at the site. See also, crispr.u-psud.fr/, a tool designed to help scientists find CRISPR targeting sites in a wide range of species and generate the appropriate crRNA sequence.

In some embodiments, one or more vectors driving expression of one
25 or more elements of a CRISPR system are introduced into a target cell such that expression of the elements of the CRISPR system direct formation of a CRISPR complex at one or more target sites. For example, a Cas enzyme, a guide sequence linked to a tracr-mate sequence, and a tracr sequence could each be operably linked to separate regulatory elements on separate vectors.
30 Alternatively, two or more of the elements expressed from the same or different regulatory elements may be combined in a single vector, with one or more additional vectors providing any components of the CRISPR system not included in the first vector. CRISPR system elements that are combined in a single vector may be arranged in any suitable orientation, such as one

element located 5' with respect to ("upstream" of) or 3' with respect to ("downstream" of) a second element. The coding sequence of one element can be located on the same or opposite strand of the coding sequence of a second element, and oriented in the same or opposite direction. In some embodiments, a single promoter drives expression of a transcript encoding a CRISPR enzyme and one or more of the guide sequence, tracr mate sequence (optionally operably linked to the guide sequence), and a tracr sequence embedded within one or more intron sequences (e.g., each in a different intron, two or more in at least one intron, or all in a single intron). In some embodiments, the CRISPR enzyme, guide sequence, tracr mate sequence, and tracr sequence are operably linked to and expressed from the same promoter.

In some embodiments, a vector includes one or more insertion sites, such as a restriction endonuclease recognition sequence (also referred to as a "cloning site"). In some embodiments, one or more insertion sites (e.g., about or more than about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more insertion sites) are located upstream and/or downstream of one or more sequence elements of one or more vectors. In some embodiments, a vector includes an insertion site upstream of a tracr mate sequence, and optionally downstream of a regulatory element operably linked to the tracr mate sequence, such that following insertion of a guide sequence into the insertion site and upon expression the guide sequence directs sequence-specific binding of a CRISPR complex to a target sequence in a eukaryotic cell. In some embodiments, a vector includes two or more insertion sites, each insertion site being located between two tracr mate sequences so as to allow insertion of a guide sequence at each site. In such an arrangement, the two or more guide sequences can include two or more copies of a single guide sequence, two or more different guide sequences, or combinations of these. When multiple different guide sequences are used, a single expression construct may be used to target CRISPR activity to multiple different, corresponding target sequences within a cell. For example, a single vector can include about or more than about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20 guide sequences. In some embodiments, about or more than about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, such

CA 03149421 2022-2-24

guide-sequence-containing vectors may be provided, and optionally delivered to a cell.

In some embodiments, a vector includes a regulatory element operably linked to an enzyme-coding sequence encoding a CRISPR enzyme, such as a Cas protein. Non-limiting examples of Cas proteins include Casl, CaslB, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7, Cas8, Cas9 (also known as Csnl and Csxl2), CaslO, Csyl, Csy2, Csy3, Csel, Cse2, Cscl, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmrl, Cmr3, Cmr4, Cmr5, Cmr6, Csbl, Csb2, Csb3, Csxl7, Csxl4, CsxlO, Csxl6, CsaX, Csx3, Csxl, Csxl5, Csfl, Csf2, Csf3, Csf4, homologues thereof, or modified versions thereof. In some embodiments, the unmodified CRISPR enzyme has DNA cleavage activity, such as Cas9. In some embodiments, the CRISPR enzyme directs cleavage of one or both strands at the location of a target sequence, such as within the target sequence and/or within the complement of the target sequence. In some embodiments, the CRISPR enzyme directs cleavage of one or both strands within about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 50, 100, 200, 500, or more base pairs from the first or last nucleotide of a target sequence.

In some embodiments, a vector encodes a CRISPR enzyme that is mutated with respect to a corresponding wild-type enzyme such that the mutated CRISPR enzyme lacks the ability to cleave one or both strands of a target polynucleotide containing a target sequence. For example, an aspartate-to-alanine substitution (D10A) in the RuvC I catalytic domain of Cas9 from *S. pyogenes* converts Cas9 from a nuclease that cleaves both strands to a nickase (cleaves a single strand). Other examples of mutations that render Cas9 a nickase include, without limitation, H840A, N854A, and N863A. As a further example, two or more catalytic domains of Cas9 (RuvC I, RuvC II, and RuvC III) can be mutated to produce a mutated Cas9 substantially lacking all DNA cleavage activity. In some embodiments, a D10A mutation is combined with one or more of H840A, N854A, or N863A mutations to produce a Cas9 enzyme substantially lacking all DNA cleavage activity. In some embodiments, a CRISPR enzyme is considered to substantially lack all DNA cleavage activity when the DNA cleavage activity of the mutated enzyme is less than about 25%, 10%, 5%>, 1%>, 0.1 %>, 0.01%, or lower with respect to its non-mutated form.

In some embodiments, an enzyme coding sequence encoding a CRISPR enzyme is codon optimized for expression in particular cells, such as eukaryotic cells. The eukaryotic cells can be those of or derived from a particular organism, such as a mammal, including but not limited to human, mouse, rat, rabbit, dog, or non-human primate. In general, codon optimization refers to a process of modifying a nucleic acid sequence for enhanced expression in the host cells of interest by replacing at least one codon (e.g., about or more than about 1, 2, 3, 4, 5, 10, 15, 20, 25, 50, or more codons) of the native sequence with codons that are more frequently or most frequently used in the genes of that host cell while maintaining the native amino acid sequence. Various species exhibit particular bias for certain codons of a particular amino acid. Codon bias (differences in codon usage between organisms) often correlates with the efficiency of translation of messenger RNA (mRNA), which is in turn believed to be dependent on, among other things, the properties of the codons being translated and the availability of particular transfer RNA (tRNA) molecules.

The predominance of selected tRNAs in a cell is generally a reflection of the codons used most frequently in peptide synthesis. Accordingly, genes can be tailored for optimal gene expression in a given organism based on codon optimization. Codon usage tables are readily available, for example, at the "Codon Usage Database", and these tables can be adapted in a number of ways. See Nakamura, Y., et al., *Nucl. Acids Res.*, 28:292 (2000). Computer algorithms for codon optimizing a particular sequence for expression in a particular host cell, for example Gene Forge (Aptagen; Jacobus, PA), are also available. In some embodiments, one or more codons (e.g., 1, 2, 3, 4, 5, 10, 15, 20, 25, 50, or more, or all codons) in a sequence encoding a CRISPR enzyme correspond to the most frequently used codon for a particular amino acid.

In some embodiments, a vector encodes a CRISPR enzyme including one or more nuclear localization sequences (NLSs). When more than one NLS is present, each may be selected independently of the others, such that a single NLS may be present in more than one copy and/or in combination with one or more other NLSs present in one or more copies. In some embodiments, an NLS is considered near the N- or C-terminus when the

CA 03149421 2022-2-24

nearest amino acid of the NLS is within about 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 40, 50, or more amino acids along the polypeptide chain from the N-or C-terminus.

In general, the one or more NLSs are of sufficient strength to drive accumulation of the CRISPR enzyme in a detectable amount in the nucleus of a eukaryotic cell. In general, strength of nuclear localization activity may derive from the number of NLSs in the CRISPR enzyme, the particular NLS(s) used, or a combination of these factors.

Detection of accumulation in the nucleus may be performed by any suitable technique. For example, a detectable marker may be fused to the CRISPR enzyme, such that location within a cell may be visualized, such as in combination with a means for detecting the location of the nucleus (e.g., a stain specific for the nucleus such as DAPI). Cell nuclei may also be isolated from cells, the contents of which may then be analyzed by any suitable process for detecting protein, such as immunohistochemistry, Western blot, or enzyme activity assay. Accumulation in the nucleus may also be determined indirectly, such as by an assay for the effect of CRISPR complex formation (e.g., assay for DNA cleavage or mutation at the target sequence, or assay for altered gene expression activity affected by CRISPR complex formation and/or CRISPR enzyme activity), as compared to a control no exposed to the CRISPR enzyme or complex, or exposed to a CRISPR enzyme lacking the one or more NLSs.

In some embodiments, one or more of the elements of CRISPR system are under the control of an inducible promoter, which can include inducible Cas, such as Cas9.

Cong, *Science*, 15:339(6121):819–823 (2013) reported heterologous expression of Cas9, tracrRNA, pre-crRNA (or Cas9 and sgRNA) can achieve targeted cleavage of mammalian chromosomes. Therefore, CRISPR system utilized in the methods disclosed herein, and thus the cargo nucleic acid, be a vector system which can include one or more vectors encoding elements of the CRISPR system which can include a first regulatory element operably linked to a CRISPR/Cas system chimeric RNA (chiRNA) polynucleotide sequence, wherein the polynucleotide sequence includes (a) a guide sequence capable of hybridizing to a target sequence in a eukaryotic cell, (b) a tracr

CA 03149421 2022-2-24

mate sequence, and (c) a tracr sequence; and a second regulatory element operably linked to an enzyme-coding sequence encoding a CRISPR enzyme which can optionally include at least one or more nuclear localization sequences. Elements (a), (b) and (c) can arranged in a 5' to 3 orientation, wherein components I and II are located on the same or different vectors of the system, wherein when transcribed, the tracr mate sequence hybridizes to the tracr sequence and the guide sequence directs sequence-specific binding of a CRISPR complex to the target sequence, and wherein the CRISPR complex can include the CRISPR enzyme complexed with (1) the guide sequence that is hybridized to the target sequence, and (2) the tracr mate sequence that is hybridized to the tracr sequence, wherein the enzyme coding sequence encoding the CRISPR enzyme further encodes a heterologous functional domain. In some embodiments, one or more of the vectors also encodes a suitable Cas enzyme, for example, Cas9. The different genetic elements can be under the control of the same or different promoters.

While the specifics can be varied in different engineered CRISPR systems, the overall methodology is similar. A practitioner interested in using CRISPR technology to target a DNA sequence (identified using one of the many available online tools) can insert a short DNA fragment containing the target sequence into a guide RNA expression plasmid. The sgRNA expression plasmid contains the target sequence (about 20 nucleotides), a form of the tracrRNA sequence (the scaffold) as well as a suitable promoter and necessary elements for proper processing in eukaryotic cells. Such vectors are commercially available (see, for example, Addgene). Many of the systems rely on custom, complementary oligos that are annealed to form a double stranded DNA and then cloned into the sgRNA expression plasmid. Co-expression of the sgRNA and the appropriate Cas enzyme from the same or separate plasmids in transfected cells results in a single or double strand break (depending of the activity of the Cas enzyme) at the desired target site.

ii. Zinc Finger Nucleases

In some embodiments, the element that induces a single or a double strand break in the target cell's genome is a nucleic acid construct or

CA 03149421 2022-2-24

constructs encoding a zinc finger nucleases (ZFNs). Thus, the nucleic acid cargo can encode a ZFN.

ZFNs are typically fusion proteins that include a DNA-binding domain derived from a zinc-finger protein linked to a cleavage domain. The most common cleavage domain is the Type IIS enzyme FokI. Fok1 catalyzes double-stranded cleavage of DNA, at 9 nucleotides from its recognition site on one strand and 13 nucleotides from its recognition site on the other. See, for example, U.S. Pat. Nos. 5,356,802; 5,436, 150 and 5,487,994; as well as Li et al. *Proc., Natl. Acad. Sci. USA* 89 (1992):4275-4279; Li et al. *Proc. Natl. Acad. Sci. USA*, 90:2764-2768 (1993); Kim et al. *Proc. Natl. Acad. Sci.* USA. 91:883-887 (1994a); Kim et al. *J. Biol. Chem.* 269:31 ,978-31,982 (1994b). One or more of these enzymes (or enzymatically functional fragments thereof) can be used as a source of cleavage domains.

The DNA-binding domain, which can, in principle, be designed to target any genomic location of interest, can be a tandem array of Cys_2His_2 zinc fingers, each of which generally recognizes three to four nucleotides in the target DNA sequence. The Cys_2His_2 domain has a general structure: Phe (sometimes Tyr)-Cys-(2 to 4 amino acids)-Cys-(3 amino acids)-Phe(sometimes Tyr)-(5 amino acids)-Leu-(2 amino acids)-His-(3 amino acids)-His. By linking together multiple fingers (the number varies: three to six fingers have been used per monomer in published studies), ZFN pairs can be designed to bind to genomic sequences 18-36 nucleotides long.

Engineering methods include, but are not limited to, rational design and various types of empirical selection methods. Rational design includes, for example, using databases including triplet (or quadruplet) nucleotide sequences and individual zinc finger amino acid sequences, in which each triplet or quadruplet nucleotide sequence is associated with one or more amino acid sequences of zinc fingers which bind the particular triplet or quadruplet sequence. See, for example, U.S. Pat. Nos. 6, 140,081; 6,453,242; 6,534,261; 6,610,512; 6,746,838; 6,866,997; 7,067,617; U.S. Published Application Nos. 2002/0165356; 2004/0197892; 2007/0154989; 2007/0213269; and International Patent Application Publication Nos. WO 98/53059 and WO 2003/016496.

CA 03149421 2022-2-24

iii. Transcription Activator-Like Effector Nucleases

In some embodiments, the element that induces a single or a double strand break in the target cell's genome is a nucleic acid construct or constructs encoding a transcription activator-like effector nuclease (TALEN). Thus, the nucleic acid cargo can encode a TALEN.

TALENs have an overall architecture similar to that of ZFNs, with the main difference that the DNA-binding domain comes from TAL effector proteins, transcription factors from plant pathogenic bacteria. The DNA-binding domain of a TALEN is a tandem array of amino acid repeats, each about 34 residues long. The repeats are very similar to each other; typically they differ principally at two positions (amino acids 12 and 13, called the repeat variable diresidue, or RVD). Each RVD specifies preferential binding to one of the four possible nucleotides, meaning that each TALEN repeat binds to a single base pair, though the NN RVD is known to bind adenines in addition to guanine. TAL effector DNA binding is mechanistically less well understood than that of zinc-finger proteins, but their seemingly simpler code could prove very beneficial for engineered-nuclease design. TALENs also cleave as dimers, have relatively long target sequences (the shortest reported so far binds 13 nucleotides per monomer) and appear to have less stringent requirements than ZFNs for the length of the spacer between binding sites. Monomeric and dimeric TALENs can include more than 10, more than 14, more than 20, or more than 24 repeats.

Methods of engineering TAL to bind to specific nucleic acids are described in Cermak, et al, *Nucl. Acids Res.* 1-11 (2011). US Published Application No. 2011/0145940, which discloses TAL effectors and methods of using them to modify DNA. Miller et al. *Nature Biotechnol* 29: 143 (2011) reported making TALENs for site-specific nuclease architecture by linking TAL truncation variants to the catalytic domain of FokI nuclease. The resulting TALENs were shown to induce gene modification in immortalized human cells. General design principles for TALE binding domains can be found in, for example, WO 2011/072246.

CA 03149421 2022-2-24

b. Donor Polynucleotides

The nuclease activity of the genome editing systems described herein cleave target DNA to produce single or double strand breaks in the target DNA. Double strand breaks can be repaired by the cell in one of two ways: non-homologous end joining, and homology- directed repair. In non-homologous end joining (NHEJ), the double-strand breaks are repaired by direct ligation of the break ends to one another. As such, no new nucleic acid material is inserted into the site, although some nucleic acid material may be lost, resulting in a deletion. In homology-directed repair (HDR), a donor polynucleotide with homology to the cleaved target DNA sequence is used as a template for repair of the cleaved target DNA sequence, resulting in the transfer of genetic information from a donor polynucleotide to the target DNA. As such, new nucleic acid material can be inserted/copied into the site.

Therefore, in some embodiments, the nucleic acid cargo is or includes a donor polynucleotide. The modifications of the target DNA due to NHEJ and/or homology-directed repair can be used to induce gene correction, gene replacement, gene tagging, transgene insertion, nucleotide deletion, gene disruption, gene mutation, etc.

Accordingly, cleavage of DNA by the genome editing composition can be used to delete nucleic acid material from a target DNA sequence by cleaving the target DNA sequence and allowing the cell to repair the sequence in the absence of an exogenously provided donor polynucleotide. Alternatively, if the genome editing composition includes a donor polynucleotide sequence that includes at least a segment with homology to the target DNA sequence, the methods can be used to add, i.e., insert or replace, nucleic acid material to a target DNA sequence (e.g., to "knock in" a nucleic acid that encodes for a protein, an siRNA, an miRNA, etc.), to add a tag (e.g., 6xHis, a fluorescent protein (e.g., a green fluorescent protein; a yellow fluorescent protein, etc.), hemagglutinin (HA), FLAG, etc.), to add a regulatory sequence to a gene (e.g., promoter, polyadenylation signal, internal ribosome entry sequence (IRES), 2A peptide, start codon, stop codon, splice signal, localization signal, etc.), to modify a nucleic acid sequence (e.g., introduce a mutation), and the like. As such, the

CA 03149421 2022-2-24

compositions can be used to modify DNA in a site- specific, i.e., "targeted", way, for example gene knock-out, gene knock-in, gene editing, gene tagging, etc. as used in, for example, gene therapy.

In applications in which it is desirable to insert a polynucleotide sequence into a target DNA sequence, a polynucleotide including a donor sequence to be inserted is also provided to the cell. By a "donor sequence" or "donor polynucleotide" or "donor oligonucleotide" it is meant a nucleic acid sequence to be inserted at the cleavage site. The donor polynucleotide typically contains sufficient homology to a genomic sequence at the cleavage site, e.g., 70%, 80%, 85%, 90%, 95%, or 100% homology with the nucleotide sequences flanking the cleavage site, e.g., within about 50 bases or less of the cleavage site, e.g., within about 30 bases, within about 15 bases, within about 10 bases, within about 5 bases, or immediately flanking the cleavage site, to support homology-directed repair between it and the genomic sequence to which it bears homology. The donor sequence is typically not identical to the genomic sequence that it replaces. Rather, the donor sequence may contain at least one or more single base changes, insertions, deletions, inversions or rearrangements with respect to the genomic sequence, so long as sufficient homology is present to support homology-directed repair. In some embodiments, the donor sequence includes a non-homologous sequence flanked by two regions of homology, such that homology-directed repair between the target DNA region and the two flanking sequences results in insertion of the non-homologous sequence at the target region.

CA 03149421 2022-2-24

2. Immunomodulation

a. CAR T Cells

The disclosed compositions and methods are particularly useful in the context of preparing lymphocytes expressing immune receptors, particularly
5 chimeric immune receptors (CIR) such as chimeric antigen receptors (CAR). Artificial immune receptors (also known and referred to herein, as chimeric T cell receptors, chimeric immunoreceptors, chimeric antigen receptors (CARs), and chimeric immune receptors (CIR)) are engineered receptors, which graft a selected specificity onto a cell. Cells modified according to the
10 discussed methods can be utilized, as discussed in more detail below, in a variety of immune therapies for treatment of cancers, infections, inflammation, and autoimmune diseases.

In particularly preferred embodiments, mRNA or DNA encoding a chimeric antigen receptor cargo is delivered to immune cells, such as
15 lymphocytes.

The cargo can be delivered to immune cells in vivo, ex vivo, or in vitro. In preferred embodiments, the cargo is mRNA, which may allow for one or more of reduced cost, ease of manufacturing, reduced side effects (e.g., cytokine storm, neurotoxicity, graft vs. host diseases, etc.). In a
20 particular embodiments, immune cells (e.g., T cells) are harvested from a subject in need of CAR T cell therapy, the compositions and methods disclosed herein are used to deliver mRNA encoding one or more CAR T cell constructs into the harvested cells, and the cells are returned to the subject. In some embodiments, the process, from initially harvesting the
25 cells to returning them to the subject, takes 1 week or less, for example, 1, 2, 3, 4, 5, 6, or 7 days. In particular embodiments, the process, from initially harvesting the cells to returning them to subject is carried in out in 1 or 2 days, or in less than 1 days, e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, or 23 hours.

30 Strategies for the design and development of chimeric antigen receptors are reviewed in Dotti, *et al.*, *Immunol Rev.* 2014 January; 257(1): . doi:10.1111/imr.12131 (35 pages), which is a specifically incorporated by reference herein in its entirety, as well as Dotti, *Molecular Therapy*, 22(5):899-890 (2014), Karlsson, *et al.*, *Cancer Gene Therapy*, 20:386-93

CA 03149421 2022-2-24

(2013), Charo, *et al.*, *Cancer Res.*, 65(5):2001-8 (2005), Jensen, *et al.*, *Immunol Rev.*, 257(1): 127–144 (2014), Eaton, *et al.*, *Gene Therapy*, 9:527-35 (2002), Barrett, *et al.*, *Annu Rev Med.*, 65: 333–347 (2014), Cartellieri, *et al.*, *Journal of Biomedicine and Biotechnology*, Volume 2010, Article ID
5 956304, 13 pages doi:10.1155/2010/956304; and U.S. Published Application Nos. 2015/0017120, 2015/0283178, 2015/0290244, 2014/0050709, and 2013/0071414.

CARs combine the antigen-binding property of monoclonal antibodies with the lytic capacity and self-renewal of T cells and have
10 several advantages over conventional T cells (Ramos and Dotti, *Expert Opin Biol Ther.*, 11:855–873 (2011), Curran, *et al.*, J Gene Med., 14:405–415 (2012), Maher, *ISRN Oncol.* 2012:278093 (2012)). CAR-T cells recognize and kill cancer cells independently of the major histocompatibility complex (MHC). Thus target cell recognition is unaffected by some of the
15 mechanisms by which tumors evade MHC-restricted T-cell recognition, for example downregulation of human leukocyte antigen (HLA) class I molecules and defective antigen processing.

Chimeric immune receptors were initially developed in the 1980s and originally included the variable (antigen binding) regions of a monoclonal
20 antibody and the constant regions of the T-cell receptor (TCR) α and β chains (Kuwana, et al., *Biochem Biophys Res Commun.*, 149:960–968 (1987)). In 1993 this design was modified to include an ectodomain, from a single chain variable fragment (scFv) from the antigen binding regions of both heavy and light chains of a monoclonal antibody, a transmembrane
25 domain, and an endodomain with a signaling domain derived from CD3-ζ. Later CARs have generally followed a similar structural design, with a co-stimulatory signaling endodomain. Accordingly, the CAR constructs utilized in the methods herein can include an antigen binding domain or ectodomain, a hinge domain, a transmembrane domain, an endodomain, and combinations
30 thereof.

In some embodiments the ectodomain is an scFv. The affinity of the scFv predicts CAR function (Hudecek, et al., *Clin Cancer Res.*, 19(12):3153-64 (2013), Chmielewski, et al., *J Immunol.*, 173:7647–7653 (2004)). Antigen binding and subsequent activation can also be modified by adding a

CA 03149421 2022-2-24

flexible linker sequence in the CAR, which allows for expression of two distinct scFvs that can recognize two different antigens (Grada, et al., *Mol Ther Nucleic Acids*, 2:e105 (2013)) (referred to as tandem CARs (TanCARs)). Tandem CARS may be more effective in killing cancers expressing low levels of each antigen individually and may also reduce the risk of tumor immune escape due by single antigen loss variants. Other ectodomains include IL13Rα2 (Kahlon, et al., *Cancer Res.*, 64:9160–9166 (2004), Brown, et al., *Clin Cancer Res.*, 18(8):2199-209 (2012), Kong, et al., *Clin Cancer Res.*, 18:5949–5960 (2012), NKG2D-ligand and CD70 receptor, peptide ligands (e.g., T1E peptide ligand), and so-called "universal ectodomains" (e.g., avidin ectodomain designed to recognize targets that have been contacted with biotinylated monoclonal antibodies, or FITC-specific scFv designed to recognize targets that have been contacted with FITC-labeled monoclonal antibodies (Zhang, et al., *Blood*, 106:1544–1551 (2005), Barber, et al., *Exp Hematol.*, 36:1318–1328 (2008), Shaffer , et al., Blood, 117:4304–4314 (2011), Davies, et al., *Mol Med.*, 18:565–576 (2012), Urbanska, et al., *Cancer Res.*, 72:1844–1852 (2012), Tamada, et al., *Clin Cancer Res.*,18:6436–6445 (2012)).

In some embodiments, the CAR includes a hinge region. While the ectodomain is important for CAR specificity, the sequence connecting the ectodomain to the transmembrane domain (the hinge region) can also influence CAR-T-cell function by producing differences in the length and flexibility of the CAR. Hinges can include, for example, a CH2CH3 hinge, or a fragment thereof, derived from an immunoglobulin such as IgG1. For example, Hudecek et al. (Hudecek, et al., *Clin Cancer Res.*, 19(12):3153-64 (2013)) compared the influence of a CH2-CH3 hinge [229 amino acids (AA)], CH3 hinge (119 AA), and short hinge (12AA) on the effector function of T cells expressing 3rd generation ROR1-specific CARs and found that T cells expressing 'short hinge' CARs had superior antitumor activity, while other investigators found that a CH2-CH3 hinge impaired epitope recognition of a 1st generation CD30-specific CAR (Hombach, et al., *Gene Ther.*, 7:1067–1075 (2000)).

Between the hinge (or ectodomain if no hinge domain) and the signaling endodomains typically lies a transmembrane domain, most

typically derived from CD3-ζ, CD4, CD8, or CD28 molecules. Like hinges, the transmembrane domain can also influence CAR-T-cell effector function.

Upon antigen recognition, CAR endodomains transmit activation and costimulatory signals to T cells. T-cell activation relies on the
5 phosphorylation of immunoreceptor tyrosine-based activation motifs (ITAMs) present in the cytoplasmic domain to the cytoplasmic CD3-ζ domain of the TCR complex (Irving, et al., *Cell*, 64:891–901 (1991)). Although the majority of CAR endomains contain an activation domain derived from CD3-ζ, others can include ITAM-containing domains such as
10 the Fc receptor for IgE-γ domain (Haynes, et al., *J Immunol.*, 166:182–187 (2001)).

The target specificity of the cell expressing a CAR is determined by the antigen recognized by the antibody/ectodomain. The disclosed compositions and methods can be used to create constructs, and cells
15 expressing the constructs, that target any antigen. In the context of immunotherapy, particularly cancer immunotherapy, numerous antigens, and suitable ectodomains for targeting them, are well known. Unlike the native TCR, the majority of scFv-based CARs recognize target antigens expressed on the cell surface rather than internal antigens that are processed and
20 presented by the cells' MHC, however, CARs have the advantage over the classical TCR that they can recognize structures other than protein epitopes, including carbohydrates and glycolipids Dotti, et al., *Immunol Rev.* 2014 January ; 257(1): . doi:10.1111/imr.12131 (35 pages) thus increasing the pool of potential target antigens. Preferred targets include antigens that are
25 only expressed on cancer cells or their surrounding stroma (Cheever, et al., *Clin Cancer Res.*,15:5323–5337 (2009)), such as the splice variant of EGFR (EGFRvIII), which is specific to glioma cells (Sampson, et al., *Semin Immunol.*, 20(5):267-75 (2008)). However, human antigens meet this requirement, and the majority of target antigens are expressed either at low
30 levels on normal cells (e.g. GD2, CAIX, HER2) and/or in a lineage restricted fashion (e.g. CD19, CD20).

Preferred targets, and CARs that target them are known in the art (see, e.g., Dotti, et al., *Immunol Rev.* 2014 January ; 257(1): . doi:10.1111/imr.12131 (35 pages). For example, CAR targets for

CA 03149421 2022-2-24

hematological malignancies include, but are not limited to, CD 19 (e.g., B-cell) (Savoldo, et al., *J Clin Invest.*, 121:1822-1826 (2011), Cooper, et al., *Blood*, 105:1622-1631 (2005); Jensen, et al., *Biol Blood Marrow Transplant* (2010), Kochenderfer, et al., *Blood*, 119:2709-2720 (2012), Brentjens, et al.,
5 *Molecular Therapy*, 17:S157 (2009), Brentjens, et al., *Nat Med.*, 9:279-286 (2003), Brentjens, et al., *Blood*, 118:4817-4828 (2011), Porter, et al., *N Engl J Med.*, 365:725-733 (2011), Kalos, et al., *Sci Transl Med.*, 3:95ra73 (2011), Brentjens, et al., *Sci Transl Med.*, 5:177ra38 (2013), Grupp, et al., *N Engl J Med* (2013)); CD20 (e.g., B-cell) (Jensen, et al., *Biol Blood Marrow*
10 *Transplant* (2010), Till, et al., *Blood*, 112:2261-2271 (2008), Wang, et al., *Hum Gene Ther.*, 18:712-725 (2007), Wang, et al., *Mol Ther.*, 9:577-586 (2004), Jensen, et al., *Biol Blood Marrow Transplant*, 4:75-83 (1998)); CD22 (e.g., B-cell) (Haso, et al., *Blood*, 121:1165-1174 (2013)); CD30 (e.g., B-cell) (Di Stasi, et al., *Blood*, 113:6392-6402 (2009), Savoldo, et al., *Blood*,
15 110:2620-2630 (2007), Hombach, et al., *Cancer Res.*, 58:1116-1119 (1998)); CD33 (e.g., Myeloid) (Finney, et al., *J Immunol.*, 161:2791-2797 (1998)); CD70 (e.g., B-cell/T-cell) (Shaffer, et al., *Blood*, 117:4304-4314 (2011)); CD123 (e.g., Myeloid) (Tettamanti, et al., *Br J Haematol.*, 161:389-401 (2013)); Kappa (e.g., B-cell) (Vera, et al., *Blood*, 108:3890-3897 (2006));
20 Lewis Y (e.g., Myeloid) (Peinert, et al., *Gene Ther.*, 17:678-686 (2010), Ritchie, et al., *Mol Ther.* (2013)); NKG2D ligands (e.g., Myeloid) (Barber, et al., *Exp Hematol.*, 36:1318-1328 (2008), Lehner, et al., *PLoS One.*, 7:e31210 (2012), Song, et al., *Hum Gene Ther.*, 24:295-305 (2013), Spear, et al., *J Immunol.* 188:6389-6398 (2012)); ROR1 (e.g., B-cell) (Hudecek, et al., *Clin*
25 *Cancer Res.* (2013)).

CAR targets for solid tumors include, but are not limited to, B7H3 (e.g., sarcoma, glioma) (Cheung, et al., *Hybrid Hybridomics*, 22:209–218 (2003)); CAIX (e.g., kidney) (Lamers, et al., *J Clin Oncol.*, 24:e20–e22. (2006)), Weijtens, et al., *Int J Cancer*, 77:181–187 (1998)); CD44 v6/v7
30 (e.g., cervical) (Hekele, et al., *Int J Cancer*, 68:232-238 (1996)), Dall, et al., *Cancer Immunol Immunother*, 54:51-60 (2005); CD171 (e.g., neuroblastoma) (Park, et al., *Mol Ther.*, 15:825-833 (2007)); CEA (e.g., colon) (Nolan, et al., *Clin Cancer Res.*, 5:3928-3941 (1999)); EGFRvIII (e.g., glioma) (Bullain, et al., *J Neurooncol.* (2009), Morgan, et al., *Hum*

CA 03149421 2022-2-24

Gene Ther., 23:1043-1053 (2012)); EGP2 (e.g., carcinomas) (Meier, et al., *Magn Reson Med.*, 65:756-763 (2011), Ren-Heidenreich, et al., *Cancer Immunol Immunother.*, 51:417-423 (2002)); EGP40 (e.g., colon) (Daly, et al., *Cancer Gene Ther.*, 7:284-291 (2000); EphA2 (e.g., glioma, lung)
5 (Chow, et al., *Mol Ther.*, 21:629-637 (2013)); ErbB2(HER2) (e.g., breast, lung, prostate, glioma) (Zhao, et al., *J Immunol.*, 183:5563-5574 (2009), Morgan, et al., *Mol Ther.*, 18:843-851 (2010), Pinthus, et al., 114:1774-1781 (2004), Teng, et al., *Hum Gene Ther.*, 15:699-708 (2004), Stancovski, et al., *J Immunol.*, 151:6577-6582 (1993), Ahmed, et al., *Mol Ther.*, 17:1779-1787
10 (2009), Ahmed, et al., *Clin Cancer Res.*, 16:474-485 (2010), Moritz, et al., *Proc Natl Acad Sci U.S.A.*, 91:4318-4322 (1994)); ErbB receptor family (e.g., breast, lung, prostate, glioma) (Davies, et al., *Mol Med.*, 18:565-576 (2012)); ErbB3/4 (e.g., breast, ovarian) (Muniappan, et al., *Cancer Gene Ther.*, 7:128-134 (2000), Altenschmidt, et al., *Clin Cancer Res.*, 2:1001-1008
15 (1996)); HLA-A1/MAGE1 (e.g., melanoma) (Willemsen, et al., *Gene Ther.*, 8:1601-1608 (2001), Willemsen, et al., *J Immunol.*, 174:7853-7858 (2005)); HLA-A2/NY-ESO-1 (e.g., sarcoma, melanoma) (Schuberth, et al., *Gene Ther.*, 20:386-395 (2013)); FR-α (e.g., ovarian) (Hwu, et al., *J Exp Med.*, 178:361-366 (1993), Kershaw, et al., *Nat Biotechnol.*, 20:1221-1227 (2002),
20 Kershaw, et al., *Clin Cancer Res.*, 12:6106-6115 (2006), Hwu, et al., *Cancer Res.*, 55:3369-3373 (1995)); FAP (e.g., cancer associated fibroblasts) (Kakarla, et al., *Mol Ther.* (2013)); FAR (e.g., rhabdomyosarcoma) (Gattenlohner, et al., *Cancer Res.*, 66:24-28 (2006)); GD2 (e.g., neuroblastoma, sarcoma, melanoma) (Pule, et al., *Nat Med.*, 14:1264-1270
25 (2008), Louis, et al., *Blood*, 118:6050-6056 (2011), Rossig, et al., *Int J Cancer.*, 94:228-236 (2001)); GD3 (e.g., melanoma, lung cancer) (Yun, et al., *Neoplasia.*, 2:449-459 (2000)); HMW-MAA (e.g., melanoma) (Burns, et al., *Cancer Res.*, 70:3027-3033 (2010)); IL11Rα (e.g., osteosarcoma) (Huang, et al., *Cancer Res.*, 72:271-281 (2012)); IL13Rα2 (e.g., glioma)
30 (Kahlon, et al., *Cancer Res.*, 64:9160-9166 (2004), Brown, et al., *Clin Cancer Res.* (2012), Kong, et al., *Clin Cancer Res.*, 18:5949-5960 (2012), Yaghoubi, et al., *Nat Clin Pract Oncol.*, 6:53-58 (2009)); Lewis Y (e.g., breast/ovarian/pancreatic) (Peinert, et al., *Gene Ther.*, 17:678-686 (2010), Westwood, et al., *Proc Natl Acad Sci U.S.A.*, 102:19051-19056 (2005),

Mezzanzanica, et al., *Cancer Gene Ther.*, 5:401-407 (1998)); Mesothelin (e.g., mesothelioma, breast, pancreas) (Lanitis, et al., *Mol Ther.*, 20:633-643 (2012), Moon, et al., *Clin Cancer Res.*, 17:4719-4730 (2011)); Mue1 (e.g., ovarian, breast, prostate) (Wilkie, et al., *J Immunol.*, 180:4901-4909 (2008));
5 NCAM (e.g., neuroblastoma, colorectal) (Gilham, et al., *J Immunother.*, 25:139-151 (2002)); NKG2D ligands (e.g., ovarian, sacoma) (Barber, et al., *Exp Hematol.*, 36:1318-1328 (2008), Lehner, et al., *PLoS One*, 7:e31210 (2012), Song, et al., *Gene Ther.*, 24:295-305 (2013), Spear, et al., *J Immunol.*, 188:6389-6398 (2012)); PSCA (e.g., prostate, pancreatic)
10 (Morgenroth, et al., *Prostate*, 67:1121-1131 (2007), Katari, et al., *HPB*, 13:643-650 (2011)); PSMA (e.g., prostate) (Maher, et al., *Nat Biotechnol.*, 20:70-75 (2002), Gong, et al., *Neoplasia.*, 1:123-127 (1999)); TAG72 (e.g., colon) (Hombach, et al., *Gastroenterology*, 113:1163-1170 (1997), McGuinness, et al., *Hum Gene Ther.*, 10:165-173 (1999)); VEGFR-2 (e.g.,
15 tumor vasculature) (*J Clin Invest.*, 120:3953-3968 (2010), Niederman, et al., *Proc Natl Acad Sci U.S.A.*, 99:7009-7014 (2002)).

b. Metabolic Stability

In some embodiments, cells' (e.g., CAR cells') metabolic stability is improved by equipping them with the capacity to make the very growth
20 factors that are limiting *in vivo*. In some embodiments, nucleic acid cargo encoding an anti-apoptotic factor such as BCL-XL is transiently delivered to cells. B-cell lymphoma-extra large (Bcl-XL, or BCL2-like 1 isoform 1) is a transmembrane protein in the mitochondria. It is a member of the Bcl-2 family of proteins, and acts as a pro-survival protein in the intrinsic apoptotic
25 pathway by preventing the release of mitochondrial contents such as cytochrome c, which would lead to caspase activation. Both amino acid and nucleic acid sequences encoding BCL-XL are known in the art and include, for example, UniProtKB - Q07817 (B2CL1_HUMAN), Isoform Bcl-X(L) (identifier: Q07817-1) (amino acid sequence); ENA|U72398|U72398.1
30 *Human Bcl-x beta (bcl-x) gene, complete cds* (genomic nucleic acid sequences); ENA|Z23115|Z23115.1 *H.sapiens bcl-XL mRNA* (mRNA/cDNA nucleic acid sequences).

In some embodiments, the nucleic cargo encodes a proliferation inducing factor such as IL-2. Both amino acid and nucleic acid sequences

CA 03149421 2022-2-24

encoding IL-2 are known in the art and include, for example, UniProtKB - P60568 (IL2_HUMAN) (amino acid sequence); ENA|X00695|X00695.1 *Human interleukin-2 (IL-2) gene and 5'-flanking region* (genic nucleic acid sequence); and ENA|V00564|V00564.1 *Human mRNA encoding interleukin-*
5 *2 (IL-2)* (mRNA/cDNA nucleic acid sequence).

However, the production of secreted IL-2 may have the unwanted side effect of also stimulating the proliferation of the lymphoma and Treg cells, and impairing the formation of memory T cells (Zhang, et al., *Nature Medicine*, 11:1238-1243 (2005)). In addition, the use of IL-2 in patients
10 treated with Tumor Infiltrating Lymphocytes (TILs) led to increased toxicity (Heemskerk, et al., *Human Gene Therapy*, 19:496-510 (2008)). To avoid this potentiality, in addition or alternative to IL-2, the nucleic acid cargo can encode a chimeric γc cytokine receptor (CγCR) such as one composed of Interleukin-7 (IL-7) tethered to IL-7Rα/CD127 that confers exogenous
15 cytokine independent, cell intrinsic, STAT5 cytokine signals (Hunter, et al., *Molecular Immunology*, 56:1-11 (2013)). The design is modular in that the IL-2Rβ/CD122 cytoplasmic chain can be exchanged for that of IL-7Rα/CD127, to enhance Shc activity. The constructs mimic wild type IL-2 signaling in human CD8+ T cells (Hunter, et al., *Molecular Immunology*,
20 56:1-11 (2013)) and should, therefore, work similarly to the IL-2 mRNA, without the unwanted to side effects.

Additionally and alternatively other antiapoptotic molecules and cytokines can be used to preserve cell viability in the native state. Exemplary factors include, but are not limited to:

25 Myeloid Cell Leukemia 1 (MCL-1) (e.g., UniProtKB - Q07820 (MCL1_HUMAN) (amino acid sequence); ENA|AF147742|AF147742.1 *Homo sapiens myeloid cell differentiation protein (MCL1) gene, promoter and complete cds* (genomic nucleic acid sequence); ENA|AF118124|AF118124.1 *Homo sapiens myeloid cell leukemia sequence*
30 *1 (MCL1) mRNA, complete cds.* (mRNA/cDNA nucleic acid sequence)) which is an antiapoptotic factor;

IL-7 (e.g., UniProtKB - P13232 (IL7_HUMAN) (amino acid sequence); ENA|EF064721|EF064721.1 *Homo sapiens interleukin 7 (IL7) gene, complete cds.* (genomic nucleic acid sequence); ENA|J04156|J04156.1

CA 03149421 2022-2-24

Human interleukin 7 (IL-7) mRNA, complete cds. (mRNA/cDNA nucleic acid sequence) which is important for T cell survival and development, and

IL-15 (e.g., UniProtKB - P40933 (IL15_HUMAN) (amino acid sequence); ENA|X91233|X91233.1 *H.sapiens IL15 gene* (genomic nucleic
5 acid sequence); ENA|U14407|U14407.1 *Human interleukin 15 (IL15) mRNA, complete cds.* (mRNA/cDNA nucleic acid sequence)) which promotes T and NK cell survival (Opferman, et al., *Nature*, 426: 671-676 (2003); Meazza, et al., *Journal of Biomedicine & Biotechnology*, 861920, doi:10.1155/2011/861920 (2011); Michaud, et al., *Journal of*
10 *Immunotherapy*, 33:382-390 (2010)). These cytokine mRNAs can be used either independently or in combination with BCL-XL, IL-2, and/or CγCR mRNA. Accordingly, in some embodiments, an mRNA encoding MCL-1, IL-7, IL-15, or a combination thereof is delivered to cells.

c. Inhibitory CAR (iCAR)

15 In some embodiments, T cell therapies are delivered to the CAR cells that have demonstrated long-term efficacy and curative potential for the treatment of some cancers, however, their use is limited by damage to non-cancerous tissues reminiscent of graft-versus-host disease after donor lymphocyte infusion. Any of the disclosed compositions and methods can be
20 used in combination with a non-specific immunosuppression (e.g., high-dose corticosteroid therapy, which exert cytostatic or cytotoxic effects on T cells, to restrain immune responses), irreversible T cell elimination (e.g., so-called suicide gene engineering strategies), or a combination thereof. However, in some preferred embodiments, off-target effects are reduced by introducing
25 into the CAR cell a construct encoding an inhibitory chimeric antigen receptor (iCAR). T cells with specificity for both tumor and off-target tissues can be restricted to tumor only by using an antigen-specific iCAR introduced into the T cells to protect the off-target tissue (Fedorov, et al., *Science Translational Medicine*, 5:215ra172 (2013)). The iCAR can include
30 a surface antigen recognition domain combined with a powerful acute inhibitory signaling domain to limit T cell responsiveness despite concurrent engagement of an activating receptor (e.g., a CAR). In preferred embodiments, the iCAR includes a single-chain variable fragment (scFv) specific for an inhibitory antigen fused to the signaling domains of an

CA 03149421 2022-2-24

immunoinhibitory receptor (e.g., CTLA-4, PD-1, LAG-3, 2B4 (CD244), BTLA (CD272), KIR, TIM-3, TGF beta receptor dominant negative analog etc.) via a transmembrane region that inhibits T cell function specifically upon antigen recognition. Once the CAR cell encounters a cell (e.g., a
5 cancer cell) that does not express the inhibitory antigen, iCAR-transduced T cells can mount a CAR-induced response against the CAR's target antigen. A DNA iCAR using an scFv specific for PSMA with the inhibitory signaling domains of either CTLA-4 or PD-1 is discussed in (Fedorov, et al., *Science Translational Medicine*, 5:215ra172 (2013)).

10 Design considerations include that observation that PD-1 was a stronger inhibitor than CTLA-4, CTLA-4 exhibited cytoplasmic localization unless a Y165G mutant was used, and that the iCAR expression level is important.

iCAR can be designed against cell type specific surface molecules.
15 In some embodiments the iCAR is designed to prevent T cells, NK cells, or other immune cell reactivity against certain tissues or cell types.

d. Reducing Endogenous Inhibitory Signaling

In some embodiments the cells are contacted with a nucleic acid cargo that reprograms the cells to prevent expression of one or more
20 antigens. For example, in some embodiments the nucleic acid cargo is or encodes an interfering RNA that prevents expression of an mRNA encoding antigens such as CTLA-4 or PD-1. This method can be used to prepare universal donor cells. RNAs used to alter the expression of allogenic antigens may be used alone or in combination with RNAs that result in de-
25 differentiation of the target cell.

Although the section above provides compositions and methods that utilized inhibitory signaling domains e.g., from CTLA-4 or PD-1 in an artificial iCAR to restrict on-target/off-tumor cytotoxicity, additionally or alternatively overall CAR cell on-tumor effector efficiency can be increased
30 by reducing the expression of endogenous inhibitory signaling in the CAR cells so that the CAR cells become resistant to the inhibitory signals of the hostile tumor microenvironment.

CTLA-4 and PD-1 inhibit T cells at different stages in activation and function. CTLA-4 regulates T cell responses to self-antigens, as knockout

CA 03149421 2022-2-24

mice spontaneously develop organ damage due to highly active, tissue-infiltrating T cells without specific antigen exposure (Tivol, et al., *Immunity*, 3:541-547 (1995); Waterhouse, et al., Science, 270:985-988 (1995)). Interestingly, conditional knockout of CTLA-4 in Treg cells recapitulates the
5 global knockout indicating that it normally functions within Tregs (Wing, et al., Science, 322:271-275 (2008)). In contrast, PD-L1 knockout mice are autoimmune prone, but do not spontaneously develop massive inflammatory cell infiltration of normal organs, indicating that it's major physiological function is to mediate negative feedback control of ongoing tissue
10 inflammation in an inducible manner (Dong, et al., Immunity, 20:327-336 (2004)). Indeed, according to the "adaptive resistance" hypothesis most tumors up-regulate PD-L1 in response to IFNγ; a key cytokine released by effector T cells including CART cells (Greenwald, et al., *Annu Rev Immunol*, 23:515-548 (2005); Carreno, et al., *Annu Rev Immunol*, 20:29-53 (2002);
15 Chen, et al., *The Journal of Clinical Investigation*, 125:3384-3391 (2015); Keir, et al., *Annu Rev Immunol*, 26:677-704 (2008); Pentcheva-Hoang, et al., *Immunological Reviews*, 229:67-87 (2009)). PD-L1 then delivers an inhibitory signal to T cells decreasing their proliferation, and cytokine and perforin production (Butte, et al., *Immunity*, 27:111-122 (2007); Chen, et al.,
20 *Immunology*, 4:336-347 (2004); Park, et al., *Blood*, 116:1291-1298 (2010); Wherry, et al., *Nat Immunol*, 12:492-499 (2011); Zou, et al., *Immunology*, 8:467-477 (2008)). In addition, reverse signaling from the T cell through B7-H1 on cancer cells induces an anti-apoptotic effect that counteracts Fas-L signaling (Azuma, et al., *Blood*, 111:3635-3643 (2008)). Azuma, et al.,
25 *Blood*, 111:3635-3643 (2008)

In light of the up-regulation of B7-H1 by cancer cells and the association of its expression with cancer progression and poor clinical outcome (Flies, et al., Journal of Immunotherapy, 30:251-260 (2007); Nishimura, et al., *Immunity*, 11:141-151 (1999); Wang, et al., *Curr Top*
30 *Microbiol Immunol*, 344:245-267 (2011)), antibodies antagonizing the PD-1 and CTLA-4 pathways have shown dramatic efficacy in solid tumors, particularly melanoma, with the combination of the two showing even more activity. The anti-CTLA-4 antibody, ipilimumab, improves overall survival in metastatic melanoma with increased T cell infiltration into tumors and

CA 03149421 2022-2-24

increased intratumoral CD8+:Treg ratios, predominantly through inhibition of Treg cells (Hamid, et al., *J Transl Med*, 9:204 (2011); Ribas, et al., *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, 15:6267-6276 (2009); Twyman-Saint, et al., *Nature*, 520:373-377 (2015)). The anti-PD-1 antibody, nivolumab, shows an overall response rate of 30-40% in metastatic melanoma (Robert, et al., *The New England Journal of Medicine*, 372:320-330 (2015); Topalian, et al., *J Clin Oncol*, 32:1020-1030 (2014)), with similar findings in early phase clinical trials for other solid tumors including metastatic renal cancer, non-small cell lung cancer and relapsed Hodgkin's Lymphoma (Ansell, et al., *The New England Journal of Medicine*, 372:311-319 (2015); Brahmer, et al., *J Clin Oncol*, 28:3167-3175 (2010); Topalian, et al., *The New England Journal of Medicine*, 366:2443-2454 (2012)). As resistance to anti-CTLA-4 antibodies in mouse melanoma models is due to up-regulation of PD-L181, the combination of both ipilimumab and nivolumab demonstrates further efficacy in both mouse models and human patients (Larkin, et al., *The New England Journal of Medicine*, 373:23-34 (2015); Spranger, et al., *J Immunother Cancer*, 2, 3, doi:10.1186/2051-1426-2-3 (2014); Yu, et al., *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, 16:6019-6028 (2010)). Given the importance of the checkpoint inhibition pathway, it is believed that PD-1/CTLA-4 inhibition will release the brake, while the chimeric antigen receptor will push on the gas pedal. Importantly, transient delivery can be utilized to only transiently release the brake so that these cells will not lead to future autoimmune disease.

i. CRISPRi

To avoid permanent genome modification and inactivation of inhibitory signals such as PD-1 and CTLA-4, the dCAS9 CRISPRi system (Larson, et al., *Nat Protoc*, 8:2180-2196 (2013)) can be utilized. Nucleic acids encoding the enzymatically-inactive dCAS9-KRAB-repression domain, fusion protein, and sgRNAs to the inhibitory signaling protein (e.g. CTLA-4, PD-1, LAG-3, 2B4 (CD244), BTLA (CD272), KIR, TIM-3, TGF beta receptor dominant negative analog, etc.) can be co-delivered into the CAR cell. One or multiple sgRNA can be utilized. sgRNA can be designed

CA 03149421 2022-2-24

to target the proximal promoter region and the coding region (nontemplate strand). An alternative approach utilizes the single-component Cpf1 CRISPR system, which is a smaller RNA to electroporate and express (Zetsche, et al., *Cell*, doi:10.1016/j.cell.2015.09.038 (2015)). Any of the
5 foregoing RNA components can also be encoded by DNA expression construct such as a vector, for example a plasmid. Thus, either RNA, DNA, or a combination thereof can serve as the nucleic acid cargo.

Although broad inhibition of CTLA-4 with ipilimumab results in autoimmune sequelae, it is believed these side–effects will be decreased by
10 restricting loss to CAR cells and transient nature of the mRNA delivery. Inhibitory function will be regained in time.

ii. Inhibitory RNAs

Nucleic acid cargo that can be delivered to cells can be or encode a functional nucleic acid or polypeptide designed to target and reduce or
15 inhibit expression or translation of an inhibitory signaling molecule mRNA; or to reduce or inhibit expression, reduce activity, or increase degradation of inhibitory signaling molecule protein. Suitable technologies include, but are not limited to, antisense molecules, siRNA, miRNA, aptamers, ribozymes, triplex forming molecules, RNAi, etc. In some embodiments, the mRNA
20 encode antagonist polypeptide that reduce inhibitory signaling.

In some embodiments, cargo that is or encodes functional RNAs suitable to reducing or silencing expression of CTLA-4, PD-1, LAG-3, 2B4 (CD244), BTLA (CD272), KIR, TIM-3, TGF beta receptor dominant negative analog, etc. alone or in combination can be delivered to cells.

25 In some embodiments, the cargo is an RNA or DNA that encodes a polypeptide that reduces bioavailability or serves as an antagonist or other negative regulator or inhibitor of CTLA-4, PD-1, LAG-3, 2B4 (CD244), BTLA (CD272), KIR, TIM-3, TGF beta receptor dominant negative analog, or another protein in an immune inhibitory pathway. The protein can be a
30 paracrine, endocrine, or autocrine. It can regulate the cell intracellularly. It can be secreted and regulate the expressing cell and/or other (e.g., neighboring) cells. It can be a transmembrane protein that regulates the expressing cell and/or other cells. The protein can be fusion protein, for example an Ig fusion protein.

CA 03149421 2022-2-24

e. Pro-apoptotic Factors

Compositions and methods for activating and reactivating apoptotic pathways are also provided. In some embodiments, the nucleic acid is or encodes a factor or agent that activates, reactivates, or otherwise enhances or
5 increases the intrinsic apoptosis pathway. Preferably the factor activates, reactivates, or otherwise enhances the intrinsic apoptosis pathway in cancer (e.g., tumor) cells, and is more preferably specific or targeted to the cancer cells.

In some embodiments, cells, following delivery of an anti-apoptotic
10 factor or pro-proliferation factor, such as those discussed above or otherwise known in the art, are more resistant or less sensitive to induced apoptosis than untreated cells. A pro-apoptotic factor can induce or increase apoptosis in, for example, untreated cells relative to the treated T cells, and is preferably selective for cancer cells. The regimen results in a two-pronged
15 attack, one cellular and one molecular, against the cancer cells.

The intrinsic apoptosis pathway can be activated, reactivated, or otherwise enhanced by targeting BCL-2 family members. BCL-2 family members are classified into three subgroups based on function and Bcl-2 Homology (BH) domains: multi-domain anti-apoptotic (e.g. BCL-2 or BCL-
20 XL), multi-domain pro-apoptotic (e.g. BAX and BAK), and BH3-only pro-apoptotic (e.g. BIM) proteins. Members of the BH3-only subgroup, such as BIM, function as death sentinels that are situated throughout the cell, poised to transmit a variety of physiological and pathologic signals of cellular injury to the core apoptotic machinery located at the mitochondrion (Danial, et al.,
25 *Cell*, 116:205-219 (2004)).

In some embodiments, the pro-apoptotic factor is a pro-apoptotic BH3-mimetic. Various pro-apoptotic BH3-mimetics can simulate the native pro-apoptotic activities of BIM and afford the ability to manipulate multiple points of the apoptotic pathway. For example, BIM SAHB (Stabilized Alpha
30 Helix of BCL-2 domains), ABT-737, and ABT-199 are pro-apoptotic BH3-mimetics designed by structural studies of the interaction between the pro-apoptotic BH3-only helical domain and the hydrophobic groove formed by the confluence of the BH1, BH2 and BH3 domains of anti-apoptotic proteins (Oltersdorf, et al., *Nature*, 435:677-681 (2005)).

CA 03149421 2022-2-24

D. Target Cells

In some embodiments, one or more particular cell types or tissue is the target of the disclosed complexes. The target cells can be in vitro, ex vivo or in a subject (i.e., in vivo). The application discussed herein can be carried out in vitro, ex vivo, or in vivo. For ex vivo application, the cells can be collected or isolated and treated in culture. Ex vivo treated cells can be administered to a subject in need thereof in therapeutically effective amount. For in vivo applications, cargo can be delivered to target cells passively, e.g., based on circulation of the composition, local delivery, etc., or can be actively targeted, e.g., with the additional a cell, tissue, organ specific targeting moiety. Thus, in some embodiments, cargo is delivered to the target cells to the exclusion of other cells. In some embodiments, cargo is delivered to target cells and non-target cells.

Target cells can be selected by the practitioner based on the desired treatment and therapy, and the intended effect of the nucleic acid cargo. For example, when the nucleic acid cargo is intended to induce cell death, the target cells may be cancer cells; when the nucleic acid cargo is intended to induce a genomic alteration, the target cells may be stem cells; when the nucleic acid cargo encodes a chimeric antigen receptor, the target cells may be immune cells.

3E10 scFv has previously been shown capable of penetrating into cell nuclei in an ENT2-dependent manner, with efficiency of nuclear uptake greatly impaired in ENT2-deficient cells (Hansen et al., *J Biol Chem* 282, 20790-20793 (2007)). ENT2 (SLC29A2) is a sodium-independent transporter that participates in the transport of purine and pyrimidine nucleosides and nucleobases, and is less sensitive to nitrobenzylmercaptopurine riboside (NBMPR) than ENT1.

In some embodiments, the target cells express ENT2 on their plasma member, their nuclear membrane, or both. Expression of ENT2 is relatively ubiquitous but varies in abundance among tissues and cell types. It has been confirmed in the brain, heart, placenta, thymus, pancreas, prostate and kidney (Griffiths, et al., *Biochem J*, 1997. 328 (Pt 3): p. 739-43, Crawford, et al., *J Biol Chem*, 1998. 273(9): p. 5288-93). Relative to other transporters, ENT2 has one of the highest mRNA expressions in skeletal muscle

CA 03149421 2022-2-24

(Baldwin, et al., *Pflugers Arch*, 2004. 447(5): p. 735-43, Govindarajan, et al., *Am J Physiol Regul Integr Comp Physiol*, 2007. 293(5): p. R1809-22). Thus, in some embodiments the target cells are brain, heart, placenta, thymus, pancreas, prostate, kidney, or skeletal muscle. Due to the high expression of
5 ENT2 by skeletal muscle, the disclosed compositions and methods may be particularly effective for delivering nucleic acid cargo to these cells, and/or higher levels of cargo may be delivered to these cells compared to other cells expressing lower levels of ENT2.

Additional, non-limiting, exemplary target cells are discussed below.

10 **1. Progenitor and Stem Cells**

The cells can be hematopoietic progenitor or stem cells. In some embodiments, particularly those related to gene editing and gene therapy the target cells are $CD34^+$ hematopoietic stem cells. Hematopoietic stem cells (HSCs), such as CD34+ cells are multipotent stem cells that give rise to all
15 the blood cell types including erythrocytes.

Stem cells can be isolated and enriched by one of skill in the art. Methods for such isolation and enrichment of $CD34^+$ and other cells are known in the art and disclosed for example in U.S. Patent Nos. 4,965,204; 4,714,680; 5,061,620; 5,643,741; 5,677,136; 5,716,827; 5,750,397 and
20 5,759,793. As used herein in the context of compositions enriched in hematopoietic progenitor and stem cells, "enriched" indicates a proportion of a desirable element (e.g. hematopoietic progenitor and stem cells) which is higher than that found in the natural source of the cells. A composition of cells may be enriched over a natural source of the cells by at least one order
25 of magnitude, preferably two or three orders, and more preferably 10, 100, 200 or 1000 orders of magnitude.

In humans, $CD34^+$ cells can be recovered from cord blood, bone marrow or from blood after cytokine mobilization effected by injecting the donor with hematopoietic growth factors such as granulocyte colony
30 stimulating factor (G-CSF), granulocyte-monocyte colony stimulating factor (GM-CSF), stem cell factor (SCF) subcutaneously or intravenously in amounts sufficient to cause movement of hematopoietic stem cells from the bone marrow space into the peripheral circulation. Initially, bone marrow cells may be obtained from any suitable source of bone marrow, e.g. tibiae,

CA 03149421 2022-2-24

femora, spine, and other bone cavities. For isolation of bone marrow, an appropriate solution may be used to flush the bone, which solution will be a balanced salt solution, conveniently supplemented with fetal calf serum or other naturally occurring factors, in conjunction with an acceptable buffer at low concentration, generally from about 5 to 25 mM. Convenient buffers include Hepes, phosphate buffers, lactate buffers, etc.

Cells can be selected by positive and negative selection techniques. Cells can be selected using commercially available antibodies which bind to hematopoietic progenitor or stem cell surface antigens, e.g. CD34, using methods known to those of skill in the art. For example, the antibodies may be conjugated to magnetic beads and immunogenic procedures utilized to recover the desired cell type. Other techniques involve the use of fluorescence activated cell sorting (FACS). The CD34 antigen, which is found on progenitor cells within the hematopoietic system of non-leukemic individuals, is expressed on a population of cells recognized by the monoclonal antibody My-10 (i.e., express the CD34 antigen) and can be used to isolate stem cell for bone marrow transplantation. My-10 deposited with the American Type Culture Collection (Rockville, Md.) as HB-8483 is commercially available as anti-HPCA 1. Additionally, negative selection of differentiated and "dedicated" cells from human bone marrow can be utilized, to select against substantially any desired cell marker. For example, progenitor or stem cells, most preferably $CD34^+$ cells, can be characterized as being any of $CD3^-$, $CD7^-$, $CD8^-$, $CD10^-$, $CD14^-$, $CD15^-$, $CD19^-$, $CD20^-$, $CD33^-$, Class II HLA^+ and Thy-1^+.

Once progenitor or stem cells have been isolated, they may be propagated by growing in any suitable medium. For example, progenitor or stem cells can be grown in conditioned medium from stromal cells, such as those that can be obtained from bone marrow or liver associated with the secretion of factors, or in medium including cell surface factors supporting the proliferation of stem cells. Stromal cells may be freed of hematopoietic cells employing appropriate monoclonal antibodies for removal of the undesired cells.

The isolated cells are contacted *ex vivo* with antibody and nucleic acid cargo complexes. Cells to which cargo has been delivered can be

CA 03149421 2022-2-24

referred to as modified cells. A solution of the complexes may simply be added to the cells in culture. It may be desirable to synchronize the cells in S-phase. Methods for synchronizing cultured cells, for example, by double thymidine block, are known in the art (Zielke, et al., *Methods Cell Biol.*, 8:107-121 (1974)).

The modified cells can be maintained or expanded in culture prior to administration to a subject. Culture conditions are generally known in the art depending on the cell type. Conditions for the maintenance of $CD34^+$ in particular have been well studied, and several suitable methods are available. A common approach to *ex vivo* multi-potential hematopoietic cell expansion is to culture purified progenitor or stem cells in the presence of early-acting cytokines such as interleukin-3. It has also been shown that inclusion, in a nutritive medium for maintaining hematopoietic progenitor cells *ex vivo*, of a combination of thrombopoietin (TPO), stem cell factor (SCF), and flt3 ligand (Flt-3L; i.e., the ligand of the flt3 gene product) was useful for expanding primitive (i.e., relatively non-differentiated) human hematopoietic progenitor cells *in vitro*, and that those cells were capable of engraftment in SCID-hu mice (Luens et al., 1998, Blood 91:1206-1215). In other known methods, cells can be maintained *ex vivo* in a nutritive medium (e.g., for minutes, hours, or 3, 6, 9, 13, or more days) including murine prolactin-like protein E (mPLP-E) or murine prolactin-like protein F (mPIP-F; collectively mPLP-E/IF) (U.S. Patent No. 6,261,841). It will be appreciated that other suitable cell culture and expansion methods can be used as well. Cells can also be grown in serum-free medium, as described in U.S. Patent No. 5,945,337.

In another embodiment, the modified hematopoietic stem cells are differentiated *ex vivo* into $CD4^+$ cells culture using specific combinations of interleukins and growth factors prior to administration to a subject using methods well known in the art. The cells may be expanded *ex vivo* in large numbers, preferably at least a 5-fold, more preferably at least a 10-fold and even more preferably at least a 20-fold expansion of cells compared to the original population of isolated hematopoietic stem cells.

In another embodiment cells, can be dedifferentiated somatic cells. Somatic cells can be reprogrammed to become pluripotent stem-like cells that can be induced to become hematopoietic progenitor cells. The

hematopoietic progenitor cells can then be treated with the compositions as described above with respect to CD34+ cells. Representative somatic cells that can be reprogrammed include, but are not limited to fibroblasts, adipocytes, and muscles cells. Hematopoietic progenitor cells from induced stem-like cells have been successfully developed in the mouse (Hanna, J. et al. *Science*, 318:1920-1923 (2007)).

To produce hematopoietic progenitor cells from induced stem-like cells, somatic cells are harvested from a host. In a preferred embodiment, the somatic cells are autologous fibroblasts. The cells are cultured and transduced with vectors encoding Oct4, Sox2, Klf4, and c-Myc transcription factors. The transduced cells are cultured and screened for embryonic stem cell (ES) morphology and ES cell markers including, but not limited to AP, SSEA1, and Nanog. The transduced ES cells are cultured and induced to produce induced stem-like cells. Cells are then screened for CD41 and c-kit markers (early hematopoietic progenitor markers) as well as markers for myeloid and erythroid differentiation.

The modified hematopoietic stem cells or modified cells including, e.g., induced hematopoietic progenitor cells, are then introduced into a subject. Delivery of the cells may be affected using various methods and includes most preferably intravenous administration by infusion as well as direct depot injection into periosteal, bone marrow and/or subcutaneous sites.

The subject receiving the modified cells may be treated for bone marrow conditioning to enhance engraftment of the cells. The recipient may be treated to enhance engraftment, using a radiation or chemotherapeutic treatment prior to the administration of the cells. Upon administration, the cells will generally require a period of time to engraft. Achieving significant engraftment of hematopoietic stem or progenitor cells typically takes weeks to months.

A high percentage of engraftment of modified hematopoietic stem cells may not be necessary to achieve significant prophylactic or therapeutic effect. It is believed that the engrafted cells will expand over time following engraftment to increase the percentage of modified cells. It is believed that in some cases, engraftment of only a small number or small percentage of

CA 03149421 2022-2-24

modified hematopoietic stem cells will be required to provide a prophylactic or therapeutic effect.

In preferred embodiments, the cells to be administered to a subject will be autologous, e.g. derived from the subject, or syngenic.

5 **2. Embryos**

In some embodiments, the compositions and methods can be used to deliver cargo to embryonic cells *in vitro*. The methods typically include contacting an embryo *in vitro* with an effective amount of antibody-cargo DNA to improve cargo transduction into the embryo. The embryo can be a 10 single cell zygote, however, treatment of male and female gametes prior to and during fertilization, and embryos having 2, 4, 8, or 16 cells and including not only zygotes, but also morulas and blastocytes, are also provided. In some embodiments, the embryo is contacted with the compositions on culture days 0-6 during or following *in vitro* fertilization.

15 The contacting can be adding the compositions to liquid media bathing the embryo. For example, the compositions can be pipetted directly into the embryo culture media, whereupon they are taken up by the embryo.

3. Immune cells

In some embodiments, the target cells are one or more types of 20 immune cells. For example, different type of cells can be utilized or otherwise targeted for immunodulation and CAR-based therapies. The preferred targeted/engineered T cells may vary depending on the tumor and goals of the adoptive therapy. Effector T cells are typically preferred because they secreted high levels of effector cytokines and were proficient 25 killers of tumor targets *in vitro* (Barrett, et al., *Annu Rev Med.*, 65: 333–347 (2014). Two complimentary lymphocyte populations with robust CAR mediated cytotoxicity are CD3-CD56+ NK cells and CD3+CD8+ T cells. Use of CD8+ T cells with CD4+ helper T cells leads to the increased presence of suppressive T-reg cells and dampened CD8+ T cell cytotoxicity. 30 Since reprogrammed CD8+ T cells are pre-activated so that they act directly on tumor cells without the need for activation in the lymph node, CD4+ T cell support is not essential.

Additionally, there is evidence that infusion of naive T cells (Rosenberg, et al. *Adv. Cancer Res.*, 25:323–388 (1977)), central memory T

CA 03149421 2022-2-24

cells (T_{CM} cells) (Berger, et al. *J. Clin. Invest.*, 118:294–305 (2008)), Th17 cells (Paulos, et al., *Sci. Transl. Med.*, 2:55–78 (2010)), and T stem memory cells (Gattinoni, et al., *Nat. Med.*, 17:1290–1297 (2012)) may all have certain advantages in certain applications due, for example, to their high replicative
5 capacity. Tumor Infiltrating Lymphocytes (TILs) also have certain advantages due to their antigen specificity and may be used in the delivery strategies disclosed herein.

Although sometime referred to as CAR cells, CAR immune, cells, and CART cells (or CAR T cells), it will be appreciated that the CAR and
10 other delivery strategies disclosed herein can also be carried out in other cell types, particularly different types of immune cells, including those discussed herein (e.g., lymphocytes, Natural Killer Cells, dendritic cells, B cells, antigen presenting cells, macrophage, etc.) and described elsewhere (see, e.g., Barrett, et al., *Annu Rev Med.*, 65: 333–347 (2014)).

15 **4. Cancer Cells and Tumors**

In some embodiments, the target cells are cancer cells. In such embodiments, methods of treatment are provided that may be useful in the context of cancer, including tumor therapy. The Examples below may indicate that DNA cargo may be delivered more generally to multiple tissues
20 and not restricted to tumors, while RNA delivery may be more selective for tumor tissue. Thus, in some embodiments, when cancer cells are the target cells, the cargo may be composed of RNA (e.g., RNA alone).

Cargos that may be delivered to cancer cells include, but are not limited to, constructs for the expression of one or more pro-apoptotic factors,
25 immunogenic factors, or tumor suppressors; gene editing compositions, inhibitory nucleic acids that target oncogenes; as well as other strategies discussed herein and elsewhere. In some embodiments, the cargo is mRNA that encodes a pro-apoptotic factor, or immunogenic factor that increases and immune response against the cells. In other embodiments, the cargo is
30 siRNA the reduces expression of an oncogene or other cancer-causing transcript.

In a mature animal, a balance usually is maintained between cell renewal and cell death in most organs and tissues. The various types of mature cells in the body have a given life span; as these cells die, new cells

CA 03149421 2022-2-24

are generated by the proliferation and differentiation of various types of stem cells. Under normal circumstances, the production of new cells is so regulated that the numbers of any particular type of cell remain constant. Occasionally, though, cells arise that are no longer responsive to normal growth-control mechanisms. These cells give rise to clones of cells that can expand to a considerable size, producing a tumor or neoplasm. A tumor that is not capable of indefinite growth and does not invade the healthy surrounding tissue extensively is benign. A tumor that continues to grow and becomes progressively invasive is malignant. The term cancer refers specifically to a malignant tumor. In addition to uncontrolled growth, malignant tumors exhibit metastasis. In this process, small clusters of cancerous cells dislodge from a tumor, invade the blood or lymphatic vessels, and are carried to other tissues, where they continue to proliferate. In this way a primary tumor at one site can give rise to a secondary tumor at another site.

The compositions and methods described herein may be useful for treating subjects having benign or malignant tumors by delaying or inhibiting the growth of a tumor in a subject, reducing the growth or size of the tumor, inhibiting or reducing metastasis of the tumor, and/or inhibiting or reducing symptoms associated with tumor development or growth.

Malignant tumors which may be treated are classified herein according to the embryonic origin of the tissue from which the tumor is derived. Carcinomas are tumors arising from endodermal or ectodermal tissues such as skin or the epithelial lining of internal organs and glands. The disclosed compositions are particularly effective in treating carcinomas. Sarcomas, which arise less frequently, are derived from mesodermal connective tissues such as bone, fat, and cartilage. The leukemias and lymphomas are malignant tumors of hematopoietic cells of the bone marrow. Leukemias proliferate as single cells, whereas lymphomas tend to grow as tumor masses. Malignant tumors may show up at numerous organs or tissues of the body to establish a cancer.

The types of cancer that can be treated with the provided compositions and methods include, but are not limited to, cancers such as vascular cancer such as multiple myeloma, adenocarcinomas and sarcomas,

CA 03149421 2022-2-24

of bone, bladder, brain, breast, cervical, colo-rectal, esophageal, kidney, liver, lung, nasopharangeal, pancreatic, prostate, skin, stomach, and uterine. In some embodiments, the disclosed compositions are used to treat multiple cancer types concurrently. The compositions can also be used to treat
5 metastases or tumors at multiple locations.

The disclosed compositions and methods can be further understood through the following numbered paragraphs.

1. A composition comprising or consisting of

(a) a 3E10 monoclonal antibody, cell-penetrating fragment
10 thereof; a monovalent, divalent, or multivalent single chain variable fragment (scFv); or a diabody; or humanized form or variant thereof, and

(b) a nucleic acid cargo comprising a nucleic acid encoding a polypeptide, a functional nucleic acid, a nucleic acid encoding a functional nucleic acid, or a combination thereof.

15 2. The composition of paragraph 1, wherein (a) comprises:

(i) the CDRs of any one of SEQ ID NO:1-6, 12, 13, 46-48, or 50-52 in combination with the CDRs of any one of SEQ ID NO:7-11, 14, or 53-58;

(ii) first, second, and third heavy chain CDRs selected from any of SEQ ID NOS:15-23, 42, or 43 in combination with first, second and third
20 light chain CDRs selected from any of SEQ ID NOS:24-30, 44, or 45;

(iii) a humanized form of (i) or (ii);

(iv) a heavy chain comprising an amino acid sequence comprising at least 85% sequence identity to any one of SEQ ID NO:1 or 2 in combination with a light chain comprising an amino acid sequence comprising at least
25 85% sequence identity to SEQ ID NO:7 or 8;

(v) a humanized form or (iv); or

(vi) a heavy chain comprising an amino acid sequence comprising at least 85% sequence identity to any one of SEQ ID NO:3-6, 46-48, or 50-52 in combination with a light chain comprising an amino acid sequence
30 comprising at least 85% sequence identity to SEQ ID NO:9-11 or 53-58.

3. The composition of paragraphs 1 or 2, wherein (a) comprises the same or different epitope specificity as monoclonal antibody 3E10, produced by ATCC Accession No. PTA 2439 hybridoma.

4. The composition of any one of paragraphs 1-3, wherein (a) is a recombinant antibody having the paratope of monoclonal antibody 3E10.

5. A composition comprising

(a) a binding protein comprising

5 (i) the CDRs of any one of SEQ ID NO:1-6, 12, 13, 46-48, or 50-52 in combination with the CDRs of any one of SEQ ID NO:7-11, 14, or 53-58;

(ii) first, second, and third heavy chain CDRs selected from SEQ ID NOS:15-23, 42, or 43 in combination with first, second and

10 third light chain CDRs selected from SEQ ID NOS:24-30, 44, or 45;

(iii) a humanized form of (i) or (ii);

(iv) a heavy chain comprising an amino acid sequence comprising at least 85% sequence identity to any one of SEQ ID NO:1 or 2 in combination with a light chain comprising an amino

15 acid sequence comprising at least 85% sequence identity to SEQ ID NO:7 or 8;

(v) a humanized form or (iv); or

(vi) a heavy chain comprising an amino acid sequence comprising at least 85% sequence identity to any one of SEQ ID

20 NO:3-6, 46-48, or 50-52 in combination with a light chain comprising an amino acid sequence comprising at least 85% sequence identity to SEQ ID NO:9-11 or 53-58, and

(b) a nucleic acid cargo comprising a nucleic acid encoding a polypeptide, a functional nucleic acid, a nucleic acid encoding a functional

25 nucleic acid, or a combination thereof.

6. The composition of any one of paragraphs 1-5, wherein (a) is bispecific.

7. The composition of paragraph 6, wherein (a) targets a cell type of interest.

30 8. The composition of any one of paragraphs 1-7, wherein (a) and (b) are non-covalently linked.

9. The composition of any one of paragraphs 1-8, wherein (a) and (b) are in a complex.

CA 03149421 2022-2-24

10. The composition of any one of paragraphs 1-9 wherein (b) comprises DNA, RNA, PNA or other modified nucleic acids, or nucleic acid analogs, or a combination thereof.

11. The composition of any one of paragraphs 1-10, wherein (b) comprises mRNA.

12. The composition of any one of paragraphs 1-11, wherein (b) comprises a vector.

13. The composition of paragraph 12, wherein the vector comprises a nucleic acid sequence encoding a polypeptide of interest operably linked to expression control sequence.

14. The composition of paragraph 13, wherein the vector is a plasmid.

15. The composition of any one of paragraphs 1-14, wherein (b) comprises a nucleic acid encoding a Cas endonuclease, a gRNA, or a combination thereof.

16. The composition of any one of paragraphs 1-15, wherein (b) comprises a nucleic acid encoding a chimeric antigen receptor polypeptide.

17. The composition of any one of paragraphs 1-16, wherein (b) comprises a functional nucleic acid.

18. The composition of any one of paragraphs 1-17, wherein (b) comprises a nucleic acid encoding a functional nucleic acid.

19. The composition of paragraphs 17 or 18, wherein the functional nucleic acid is antisense molecules, siRNA, miRNA, aptamers, ribozymes, RNAi, or external guide sequences.

20. The composition of any one of paragraphs 1-19, wherein (b) comprises a plurality of a single nucleic acid molecules.

21. The composition of any one of paragraphs 1-19, wherein (b) comprises a plurality of 2, 3, 4, 5, 6, 7, 8, 9, 10, or more different nucleic acid molecules.

22. The composition of any one of paragraphs 1-21, wherein (b) comprises or consists of nucleic acid molecules between about 1 and 25,000 nucleobases in length.

CA 03149421 2022-2-24

23. The composition of any one of paragraphs 1-22, wherein (b) comprises or consists of single stranded nucleic acids, double stranded nucleic acids, or a combination thereof.

24. The composition of any one of paragraphs 1-23, further comprising carrier DNA.

25. The composition of paragraph 24, wherein the carrier DNA is non-coding DNA.

26. The composition of paragraphs 24 or 25, wherein (b) is composed of RNA.

27. A pharmaceutical composition comprising the composition of any one of paragraphs 1-26 and a pharmaceutically acceptable excipient.

28. The composition of paragraph 27 further comprising polymeric nanoparticles encapsulating a complex of (a) and (b).

29. The composition of paragraph 28, wherein a targeting moiety, a cell penetrating peptide, or a combination thereof is associated with, linked, conjugated, or otherwise attached directly or indirectly to the nanoparticle.

30. A method of delivering a nucleic acid cargo to a cell comprising contacting the cell with an effective amount of the composition of any one of paragraphs 1-29.

31. The method of paragraph 30, wherein the contacting occurs ex vivo.

32. The method of paragraph 31, wherein the cells are hematopoietic stem cells, or T cells.

33. The method of any one of paragraphs 30-32, further comprising administering the cells to a subject in need thereof.

34. The method of paragraph 33, wherein the cells are administered to the subject in an effective amount to treat one or more symptoms of a disease or disorder.

35. The method of paragraph 30 wherein the contacting occurs in vivo following administration to a subject in need thereof.

36. The method of any one of paragraphs 33-35, wherein the subject has a disease or disorder.

37. The method of paragraph 36, wherein the disease or disorder is a genetic disorder, cancer, or an infection or infectious disease.

CA 03149421 2022-2-24

38. The method of paragraphs 36 or 37, wherein (b) is delivered into cells of the subject in an effective amount to reduce one or more symptoms of the disease or disorder in the subject.

39. A method of making the composition of any one of paragraphs 1-29 comprising incubating and/or mixing of (a) and (b) for an effective amount of time and at a suitable temperature to form complexes of (a) and (b), prior to contact with cells.

40. A method of making the composition of any one of paragraphs 1-29, comprising incubating and/or mixing of (a) and (b) for between about 1 min and about 30 min, about 10 min and about 20 min, or about 15 min, optionally at room temperature or 37 degrees Celsius.

41. A composition or method of any one of paragraphs 1-40 wherein 3E10 monoclonal antibody, cell-penetrating fragment thereof; a monovalent, divalent, or multivalent single chain variable fragment (scFv); or a diabody; or humanized form or variant thereof comprising the nucleic acid binding pocket of SEQ ID NOS:92 or 93, or a variant thereof with same or improved ability to bind to a nucleic acid.

42. A composition or method of any one of paragraphs 1-41 wherein the amino acid residue corresponding with D31 or N31 of a heavy chain amino acid sequence or a CDR thereof is substituted with R.

43. A composition or method of any one of paragraphs 1-42 wherein the amino acid residue corresponding with D31 or N31 of a heavy chain amino acid sequence or a CDR thereof is substituted with L.

44. A binding protein comprising

(i) a variant of CDRs of any one of SEQ ID NO:1-6, 12, 13, 46-48, or 50-52 in combination with the CDRs of any one of SEQ ID NO:7-11, 14, or 53-58;

(ii) a variant of the first heavy chain CDR, in combination with the second, and third heavy chain CDRs selected from SEQ ID NOS:15-23, 42, or 43 in combination with first, second and third light chain CDRs selected from SEQ ID NOS:24-30, 44, or 45;

(iii) a humanized form of (i) or (ii);

(iv) a heavy chain comprising an amino acid sequence comprising at least 85% sequence identity to any one of SEQ ID

CA 03149421 2022-2-24

NO:1 or 2 in combination with a light chain comprising an amino acid sequence comprising at least 85% sequence identity to SEQ ID NO:7 or 8;

(v) a humanized form or (iv); or

5 (vi) a heavy chain comprising an amino acid sequence comprising at least 85% sequence identity to any one of SEQ ID NO:3-6, 46-48, or 50-52 in combination with a light chain comprising an amino acid sequence comprising at least 85% sequence identity to SEQ ID NO:9-11 or 53-58,

10 wherein the amino acid residue corresponding with D31 or N31 is substituted with R or L.

45. The binding protein of paragraph 44, comprising the nucleic acid binding pocket of SEQ ID NOS:92 or 93, or a variant thereof with same or improved ability to bind to a nucleic acid.

15 **Examples**

With respect to the experiments below, standard 3E10 sequence was used except wherein noted to be the D31N variant (e.g., Example 4). Both standard 3E10 and the D31N variant were used as full length antibodies.

Example 1: 3E10 increases cellular uptake of PNA after 1 hour.

20 **Materials and Methods**

PNA alone (1 nmole) (MW=9984.39; 29 nucleotides in length), or PNA complexed with 3E10 (0.75 mg), was mixed at room temperature for 5 minutes. 200,000 K562 cells were then added to the suspension of 3E10, or PNA alone, in serum free media. Additional serum free media was added to 25 a final volume of 500 ul. Following incubation with cells at 37°C for 1 hr, the cells were centrifuged and washed three times with PBS prior to analysis by flow cytometry. The PNA was labeled by attachment to the fluorescent dye, tetramethylrhodamine (TAMRA).

Results

30 The results are illustrated in flow cytometry dot plots (Figure 1A-1C). % uptake was quantified (Figure 1D).

The results show increased uptake of PNA when mixed with 3E10.

CA 03149421 2022-2-24

Example 2: 3E10 increases cellular uptake of PNA after 24 hours.

Materials and Methods

PNA alone (1 nmole) (MW=9984.39; 29 nucleotides in length), or PNA complexed with 3E10 (0.75 mg), was mixed at room temperature for 5 minutes. 200,000 K562 cells were then added to the suspension of 3E10, or PNA alone, in serum free media. Additional serum free media was added to a final volume of 500 ul. Following incubation with cells at 37°C for 24 hrs, the cells were centrifuged and washed three times with PBS prior to analysis by flow cytometry.

20,000 U2OS cells were seeded in 8-well chamber slides and allowed to adhere for 24 hours. Cells were subsequently treated with PNA alone (1 nmole), or PNA complexed with 3E10 (10 uM). Following incubation at 37°C for 24 hrs, PNA or PNA mixed with 3E10 was washed with PBS prior to fixation and nuclear staining. PNA uptake was subsequently quantified by flow cytometry and imaged using fluorescent microscopy. The PNA was labeled by attachment to the fluorescent dye, tetramethylrhodamine (TAMRA).

Results

The results are illustrated in flow cytometry dot plots (Figure 2A-2C). % uptake was quantified (Figure 2D).

The results show increased uptake of PNA when mixed with 3E10.

Fluorescent microscopy showed co-localization of nuclear DNA (DAPI in blue) and PNA (Tamra in red) evident by the production of a distinct pink hue.

Example 3: 3E10 increases cellular uptake of siRNA after 24 hours.

Materials and Methods

Labeled siRNA (via attachment to fluorescein amidite, FAM) (1 nmole), or siRNA complexed with 3E10 (0.75 mg), was mixed at room temperature for 5 minutes. 200,000 K562 cells were then added to the suspension of 3E10, or siRNA alone, in serum free media. Additional serum free media was added to a final volume of 500 ul. Following incubation with cells at 37°C for 24 hrs, the cells were centrifuged and washed three times with PBS prior to analysis by flow cytometry.

CA 03149421 2022-2-24

Results

The results are illustrated in flow cytometry dot plots (Figure 3A-3C). % uptake was quantified (Figure 3D).

The results show increased cell uptake of siRNA when mixed with 3E10.

Example 4: 3E10 increases cellular uptake of mRNA after 24 hours.

Materials and Methods

Labeled mRNA (by attachment to cyanine 5, Cy5) (2 ug) alone or labeled mRNA complexed with 3E10 (2.5, 5, and 10 uM), were mixed at room temperature for 5 minutes. The suspensions of 3E10 plus mRNA, or mRNA alone, were added to 200,000 K562 cells in serum free media. Additional serum free media was added to a final volume of 500 ul. Following incubation with cells at 37°C for 24 hrs, the cells were centrifuged and washed three times with PBS prior to analysis by flow cytometry.

Results

The results are illustrated in flow cytometry dot plots (Figure 4A-4H). % uptake was quantified (Figure 4I).

The results show increased uptake of mRNA when mixed with 3E10. Note that delivery of mRNA by the D31N variant of 3E10 resulted in the highest levels of mRNA cell uptake.

Fluorescent microscopy showed functional GFP expression in U2OS cells after translation of the same Cy5 labeled mRNA, which encodes for a green fluorescent protein (GFP) reporter.

Example 5: 3E10 increases cellular uptake of mRNA after 1 hours.

Materials and Methods

Labeled mRNA (Cy5) (2 ug) or labeled mRNA complexed with the D31N variant of 3E10 (0.1-10 uM) were mixed at room temperature for 5 minutes. The suspensions of 3E10 plus mRNA, or mRNA alone, were added to 200,000 K562 cells in serum free media. Additional serum free media was added to a final volume of 500 ul. Following incubation with cells at 37°C for 1 hr, the cells were centrifuged and washed three times with PBS prior to analysis by flow cytometry.

CA 03149421 2022-2-24

Results

The results are illustrated in flow cytometry dot plots (Figure 5A-5H). % uptake was quantified (Figure 5I).

Example 6: 3E10 increases cellular uptake of plasmid DNA

5 **Materials and Methods**

GFP reporter plasmid DNA (250 ug) was complexed with 3E10 (10 uM) at room temperature for 5 minutes. The suspension of 3E10 plus plasmid DNA, or plasmid DNA alone, were added to 200,000 K562 cells in serum free media. Additional serum free media was added to a final volume
10 of 500 ul. Following incubation with cells at 37°C for 24 hrs, the cells were centrifuged and washed three times with PBS. 72 hours after the initial treatment, cells were imaged and analyzed for GFP expression.

Results

Results indicate that GFP plasmid was robustly taken up by cells
15 when 3E10 was combined with the plasmid DNA, as measured by green fluorescence, indicating uptake and functional expression of the GFP construct. No uptake or green fluorescence was seen when plasmid DNA alone was used. (Figure 6).

Example 7: 3E10 mediates mRNA delivery *in vivo*

20 **Materials and Methods**

10 ug of mRNA encoding GFP was mixed with 0.1 mg of 3E10 for 15 minutes at room temperature. mRNA complexed to 3E10 was injected systemically to BALB/c mice bearing EMT6 flank tumors measuring 100 mm^3. 20 hours after treatment, tumors were harvested and analyzed for
25 mRNA expression (GFP) using IVIS imaging.

Results

3E10-mediated delivery of mRNA resulted in significantly higher levels of GFP expression in the tumor compared to freely injected mRNA, which did not yield any GFP expression in the tumor. There was no
30 detectable expression of GFP in any of the normal tissues examined with either treatment, including liver, spleen, heart, and kidney. The results indicate robust delivery of mRNA into tumors, with functional translation and expression.

Example 8: 3E10 mediates siRNA delivery *in vivo*

Materials and Methods

40 ug of fluorescently labeled siRNA was mixed with increased doses of 3E10 (0.25, 0.5, and 1 mg) for 15 minutes at room temperature. siRNA complexed to 3E10 was injected systemically to BALB/c mice bearing EMT6 flank tumors measuring 100 mm^3. 20 hours after treatment, tumors were harvested and analyzed for siRNA delivery using IVIS imaging.

40 ug of fluorescently labeled siRNA was mixed with 1 mg 3E10 or 0.1 mg of the D31N variant of 3E10 for 15 minutes at room temperature. siRNA complexed to 3E10 was injected systemically to BALB/c mice bearing EMT6 flank tumors measuring 100 mm^3. 20 hours after treatment, tumors were harvested and analyzed for siRNA delivery using IVIS imaging.

Results

As shown in Figure 7A, increasing doses of 3E10 result in higher accumulation of siRNA in tumors.

As shown in Figure 7B, a tenfold lower dose of D31N 3E10 resulted in similar levels of siRNA delivery as 3E10.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of skill in the art to which the disclosed invention belongs. Publications cited herein and the materials for which they are cited are specifically incorporated by reference.

Example 9: Carrier DNA enhances mRNA to non-Tumor Tissue

Materials and Methods

2 ug of fluorescently labeled mRNA was mixed with 20 ug of 3E10-D31N with or without carrier DNA (5 ug) for 15 minutes at room temperature. mRNA complexed to 3E10 was injected to fetuses at E15.5. 24-48 hours after treatment, fetuses were harvested and analyzed for mRNA delivery using IVIS imaging.

Results

Without carrier DNA, 3E10-D31N complexed to mRNA was rapidly cleared from fetuses at 24 hours. The addition of carrier DNA, however, resulted in detectable mRNA signal in multiple tissues of the fetus at 48 hours.

CA 03149421 2022-2-24

The Examples above may indicate that DNA cargo delivery may be more general to multiple tissues and not restricted to tumors, while RNA delivery may be more selective for tumor tissue.

Example 10: 3E10 (D31N) Complexed with mRNA and Carrier DNA Results in Sustained Levels Protein Expression

Materials and Methods

10 ug of luciferase mRNA and 10 ug of single stranded carrier DNA (60 nts) was mixed with 100 ug of 3E10 (WT) or 3E10 (D31N) for 15 minutes at room temperature. mRNA complexed to 3E10 was injected intramuscularly (IM) in the right quadricep of each mouse. Luciferase expression was monitored over 6 days.

Results

As seen in Figure 8, administration of 3E10 (D31N) complexed with mRNA and carrier DNA resulted in sustained levels of luciferase expression, while 3E10 (WT) complexed to mRNA and carrier DNA failed to produce any appreciable signal above background.

Example 11: Distribution of IV injected 3E10 in vivo.

Distribution of IV injected 3E10 to muscle was investigated. Mice were injected intravenously with 200 µg of 3E10, WT or D31N, labeled with VivoTag680 (Perkin Elmer). Four hours after injection, muscle was harvested and imaged by IVIS (Perkin Elmer) (Figures 9A and 9B). Quantification of IVIS image demonstrates that 3E10-D31N achieves higher distribution to muscle when compared to 3E10-WT (Figure 9C).

Dose-dependent biodistribution of 3E10-D31N to tissues was investigated. Mice were injected intravenously with 100 µg or 200 µg of 3E10-D31N labeled with VivoTag680 (Perkin Elmer). 24 hours after injection, tissues were harvested and imaged by IVIS (Perkin Elmer). Quantification of tissue distribution demonstrated a dose-dependent, two-fold increase in muscle accumulation without a commensurate increase in multiple tissues including liver (Figure 10).

Distribution of 3E10 to tumors. Mice bearing flank syngeneic colon tumors (CT26) were injected intravenously with 200 µg of 3E10, WT or D31N, labeled with VivoTag680 (Perkin Elmer). 24 hours after injection, tumors were harvested and imaged by IVIS (Perkin Elmer) (Figure 11A-11B).

CA 03149421 2022-2-24

Quantification of tumor distribution demonstrated that 3E10-D31N had higher accumulation in tumors when compared to 3E10-WT (Figure 11C).

Distribution of ssDNA non-covalently associated with 3E10 was investigated. Mice bearing flank syngeneic colon tumors (CT26) were injected intravenously with 200 ug of 3E10, WT or D31N, mixed with 40 ug of labeled ssDNA (IR680). 24 hours after injection, tumors were harvested and imaged by IVIS (Perkin Elmer) (Figure 12A-12C). Quantification of tissue distribution demonstrated that delivery of ssDNA by 3E10-D31N resulted in higher tumor accumulation when compared to 3E10-WT (Figure 12D).

Example 12: 3E10-mediates delivery of RIG-I ligand, and stimulation of RIG-I activity.

Materials and Methods

RIG-I reporter cells (HEK-Lucia RIG-I, Invivogen) were seeded at 50,000 cells per well and treated with RIG-I ligands (1ug) or ligands complexed to 3E10-D31N (20 ug). This assay uses a cell line with a luciferase reporter that is activated when there is induction of interferons.

Results

In all cases, RIG-I ligands alone did not stimulate IFN-γ secretion. Delivery of RIG-ligands with 3E10-D31N, however, stimulated IFN-γ secretion above controls, with the highest secretion observed for poly (I:C), both low and high molecular weight (LMW and HMW).

Example 13: Molecular modeling of 3E10 and engineered variants thereof.

WT HEAVY CHAIN scFv SEQUENCE

E VQLVESGGGL VKPGGSRKLS CAASGFTFSD YGMHWVRQAP EKGLEWVAYI SSGSSTIYYA DTVKGRFTIS RDNAKNTLFL QMTSLRSEDT AMYYCARRGL LLDYWGQGTT LTVS (SEQ ID NO:92)

LIGHT CHAIN scFv SEQUENCE

D IVLTQSPASL AVSLGQRATI SCRASKSVST SSYSYMHWYQ QKPGQPPKLL IKYASYLESG VPARFSGSGS GTDFTLNIHP VEEEDAATYY CQHSREFPWT FGGGTKLEIK RADAAPGGGG SGGGGSGGGGS (SEQ ID NO:93)

Molecular modeling of 3E10 (Pymol) revealed a putative Nucleic Acid Binding pocket (NAB1) (Figures 14A-14B). Mutation of aspartic acid

CA 03149421 2022-2-24

at residue 31 of CDR1 to asparagine increased the cationic charge of this residue and enhanced nucleic acid binding and delivery in vivo (3E10-D31N).

Mutation of aspartic acid at residue 31 of CDR1 to arginine (3E10-
5 D31R), further expanded the cationic charge while mutation to lysine (3E10-D31K) changed charge orientation (Figure 14A).

NAB1 amino acids predicted from molecular modeling have been underlined in the heavy and light chain sequences above. Figure 14B is an illustration showing molecular modeling of 3E10-scFv (Pymol) with NAB1
10 amino acid residues illustrated with punctate dots.

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

15

We claim:

1. A composition comprising or consisting of

(a) a 3E10 monoclonal antibody, cell-penetrating fragment thereof; a monovalent, divalent, or multivalent single chain variable fragment (scFv); or a diabody; or humanized form or variant thereof, and

(b) a nucleic acid cargo comprising a nucleic acid encoding a polypeptide, a functional nucleic acid, a nucleic acid encoding a functional nucleic acid, or a combination thereof.

2. The composition of claim 1, wherein (a) comprises:

(i) the CDRs of any one of SEQ ID NO:1-6, 12, 13, 46-48, or 50-52 in combination with the CDRs of any one of SEQ ID NO:7-11, 14, or 53-58;

(ii) first, second, and third heavy chain CDRs selected from any of SEQ ID NOS:15-23, 42, or 43 in combination with first, second and third light chain CDRs selected from any of SEQ ID NOS:24-30, 44, or 45;

(iii) a humanized form of (i) or (ii);

(iv) a heavy chain comprising an amino acid sequence comprising at least 85% sequence identity to any one of SEQ ID NO:1 or 2 in combination with a light chain comprising an amino acid sequence comprising at least 85% sequence identity to SEQ ID NO:7 or 8;

(v) a humanized form or (iv); or

(vi) a heavy chain comprising an amino acid sequence comprising at least 85% sequence identity to any one of SEQ ID NO:3-6, 46-48, or 50-52 in combination with a light chain comprising an amino acid sequence comprising at least 85% sequence identity to SEQ ID NO:9-11 or 53-58.

3. The composition of claims 1 or 2, wherein (a) comprises the same or different epitope specificity as monoclonal antibody 3E10, produced by ATCC Accession No. PTA 2439 hybridoma.

4. The composition of any one of claims 1-3, wherein (a) is a recombinant antibody having the paratope of monoclonal antibody 3E10.

5. A composition comprising

(a) a binding protein comprising

(i) the CDRs of any one of SEQ ID NO:1-6, 12, 13, 46-48, or 50-52 in combination with the CDRs of any one of SEQ ID NO:7-11, 14, or 53-58;

(ii) first, second, and third heavy chain CDRs selected from SEQ ID NOS:15-23, 42, or 43 in combination with first, second and third light chain CDRs selected from SEQ ID NOS:24-30, 44, or 45;

(iii) a humanized form of (i) or (ii);

(iv) a heavy chain comprising an amino acid sequence comprising at least 85% sequence identity to any one of SEQ ID NO:1 or 2 in combination with a light chain comprising an amino acid sequence comprising at least 85% sequence identity to SEQ ID NO:7 or 8;

(v) a humanized form or (iv); or

(vi) a heavy chain comprising an amino acid sequence comprising at least 85% sequence identity to any one of SEQ ID NO:3-6, 46-48, or 50-52 in combination with a light chain comprising an amino acid sequence comprising at least 85% sequence identity to SEQ ID NO:9-11 or 53-58, and

(b) a nucleic acid cargo comprising a nucleic acid encoding a polypeptide, a functional nucleic acid, a nucleic acid encoding a functional nucleic acid, or a combination thereof.

6. The composition of any one of claims 1-5, wherein (a) is bispecific.

7. The composition of claim 6, wherein (a) targets a cell type of interest.

8. The composition of any one of claims 1-7, wherein (a) and (b) are non-covalently linked.

9. The composition of any one of claims 1-8, wherein (a) and (b) are in a complex.

10. The composition of any one of claims 1-9 wherein (b) comprises DNA, RNA, PNA or other modified nucleic acids, or nucleic acid analogs, or a combination thereof.

CA 03149421 2022-2-24

11. The composition of any one of claims 1-10, wherein (b) comprises mRNA.

12. The composition of any one of claims 1-11, wherein (b) comprises a vector.

13. The composition of claim 12, wherein the vector comprises a nucleic acid sequence encoding a polypeptide of interest operably linked to expression control sequence.

14. The composition of claim 13, wherein the vector is a plasmid.

15. The composition of any one of claims 1-14, wherein (b) comprises a nucleic acid encoding a Cas endonuclease, a gRNA, or a combination thereof.

16. The composition of any one of claims 1-15, wherein (b) comprises a nucleic acid encoding a chimeric antigen receptor polypeptide.

17. The composition of any one of claims 1-16, wherein (b) comprises a functional nucleic acid.

18. The composition of any one of claims 1-17, wherein (b) comprises a nucleic acid encoding a functional nucleic acid.

19. The composition of claims 17 or 18, wherein the functional nucleic acid is antisense molecules, siRNA, miRNA, aptamers, ribozymes, RNAi, or external guide sequences.

20. The composition of any one of claims 1-19, wherein (b) comprises a plurality of a single nucleic acid molecules.

21. The composition of any one of claims 1-19, wherein (b) comprises a plurality of 2, 3, 4, 5, 6, 7, 8, 9, 10, or more different nucleic acid molecules.

22. The composition of any one of claims 1-21, wherein (b) comprises or consists of nucleic acid molecules between about 1 and 25,000 nucleobases in length.

23. The composition of any one of claims 1-22, wherein (b) comprises or consists of single stranded nucleic acids, double stranded nucleic acids, or a combination thereof.

24. The composition of any one of claims 1-23, further comprising carrier DNA.

25. The composition of claim 24, wherein the carrier DNA is non-coding DNA.

26. The composition of claims 24 or 25, wherein (b) is composed of RNA.

27. A pharmaceutical composition comprising the composition of any one of claims 1-26 and a pharmaceutically acceptable excipient.

28. The composition of claim 27 further comprising polymeric nanoparticles encapsulating a complex of (a) and (b).

29. The composition of claim 28, wherein a targeting moiety, a cell penetrating peptide, or a combination thereof is associated with, linked, conjugated, or otherwise attached directly or indirectly to the nanoparticle.

30. A method of delivering a nucleic acid cargo to a cell comprising contacting the cell with an effective amount of the composition of any one of claims 1-29.

31. The method of claim 30, wherein the contacting occurs ex vivo.

32. The method of claim 31, wherein the cells are hematopoietic stem cells, or T cells.

33. The method of any one of claims 30-32, further comprising administering the cells to a subject in need thereof.

34. The method of claim 33, wherein the cells are administered to the subject in an effective amount to treat one or more symptoms of a disease or disorder.

35. The method of claim 30 wherein the contacting occurs in vivo following administration to a subject in need thereof.

36. The method of any one of claims 33-35, wherein the subject has a disease or disorder.

37. The method of claim 36, wherein the disease or disorder is a genetic disorder, cancer, or an infection or infectious disease.

38. The method of claims 36 or 37, wherein (b) is delivered into cells of the subject in an effective amount to reduce one or more symptoms of the disease or disorder in the subject.

CA 03149421 2022-2-24

39. A method of making the composition of any one of claims 1-29 comprising incubating and/or mixing of (a) and (b) for an effective amount of time and at a suitable temperature to form complexes of (a) and (b), prior to contact with cells.

40. A method of making the composition of any one of claims 1-29, comprising incubating and/or mixing of (a) and (b) for between about 1 min and about 30 min, about 10 min and about 20 min, or about 15 min, optionally at room temperature or 37 degrees Celsius.

41. A composition or method of any one of claims 1-40 wherein 3E10 monoclonal antibody, cell-penetrating fragment thereof; a monovalent, divalent, or multivalent single chain variable fragment (scFv); or a diabody; or humanized form or variant thereof comprising the nucleic acid binding pocket of SEQ ID NOS:92 or 93, or a variant thereof with same or improved ability to bind to a nucleic acid.

42. A composition or method of any one of claims 1-41 wherein the amino acid residue corresponding with D31 or N31 of a heavy chain amino acid sequence or a CDR thereof is substituted with R.

43. A composition or method of any one of claims 1-42 wherein the amino acid residue corresponding with D31 or N31 of a heavy chain amino acid sequence or a CDR thereof is substituted with L.

44. A binding protein comprising

(i) a variant of CDRs of any one of SEQ ID NO:1-6, 12, 13, 46-48, or 50-52 in combination with the CDRs of any one of SEQ ID NO:7-11, 14, or 53-58;

(ii) a variant of the first heavy chain CDR, in combination with the second, and third heavy chain CDRs selected from SEQ ID NOS:15-23, 42, or 43 in combination with first, second and third light chain CDRs selected from SEQ ID NOS:24-30, 44, or 45;

(iii) a humanized form of (i) or (ii);

(iv) a heavy chain comprising an amino acid sequence comprising at least 85% sequence identity to any one of SEQ ID NO:1 or 2 in combination with a light chain comprising an amino

acid sequence comprising at least 85% sequence identity to SEQ ID NO:7 or 8;

(v) a humanized form or (iv); or

(vi) a heavy chain comprising an amino acid sequence comprising at least 85% sequence identity to any one of SEQ ID NO:3-6, 46-48, or 50-52 in combination with a light chain comprising an amino acid sequence comprising at least 85% sequence identity to SEQ ID NO:9-11 or 53-58,

wherein the amino acid residue corresponding with D31 or N31 is substituted with R or L.

45. The binding protein of claim 44, comprising the nucleic acid binding pocket of SEQ ID NOS:92 or 93, or a variant thereof with same or improved ability to bind to a nucleic acid.

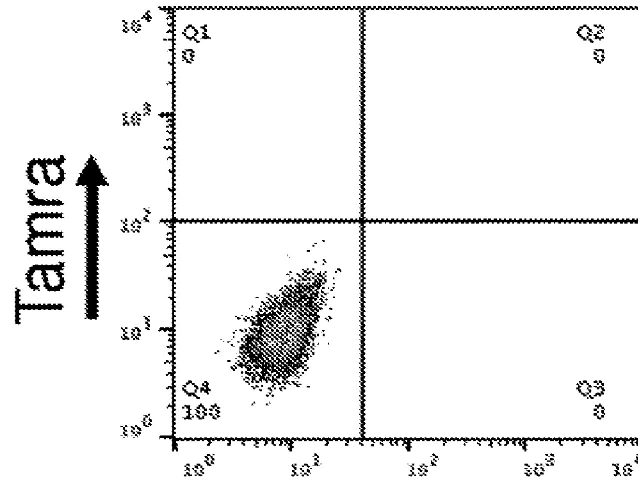


FIG. 1A

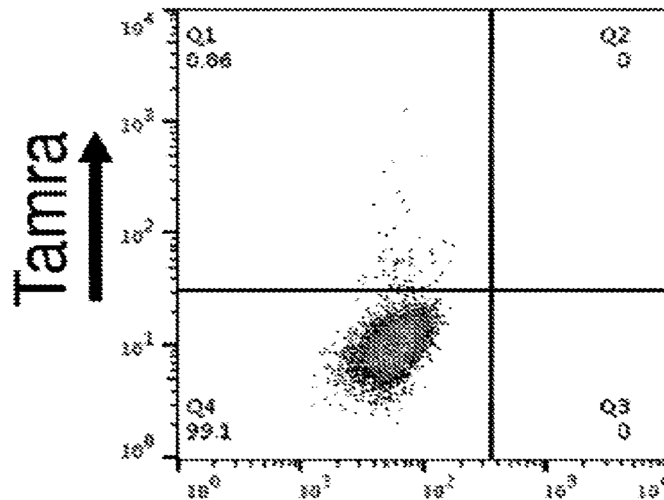


FIG. 1B

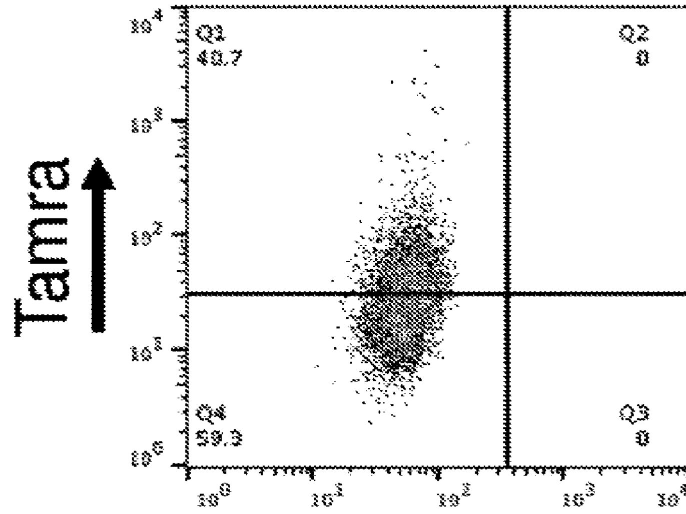


FIG. 1C

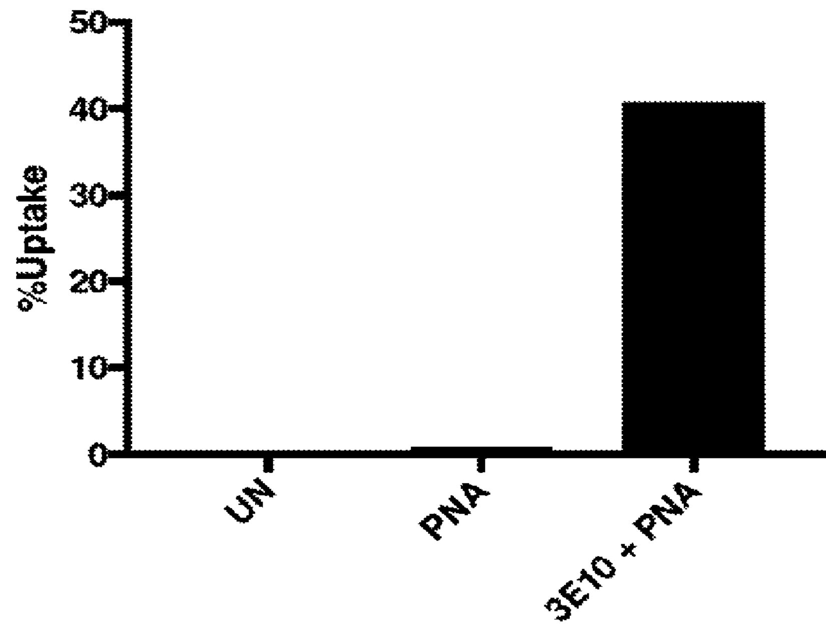


FIG. 1D

Untreated

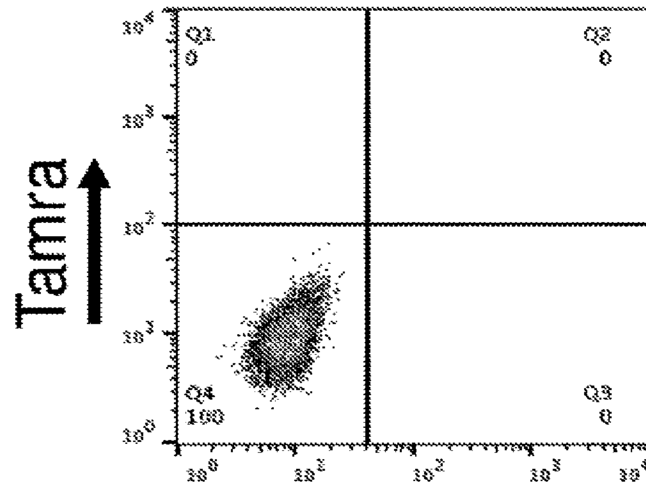


FIG. 2A

PNA Only

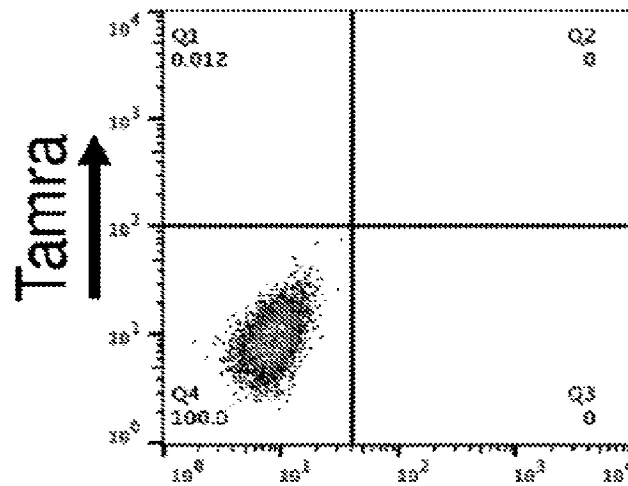


FIG. 2B

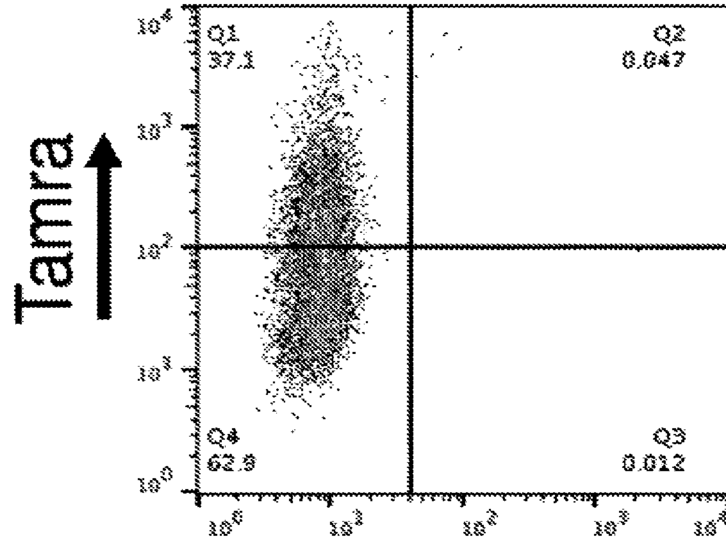


FIG. 2C

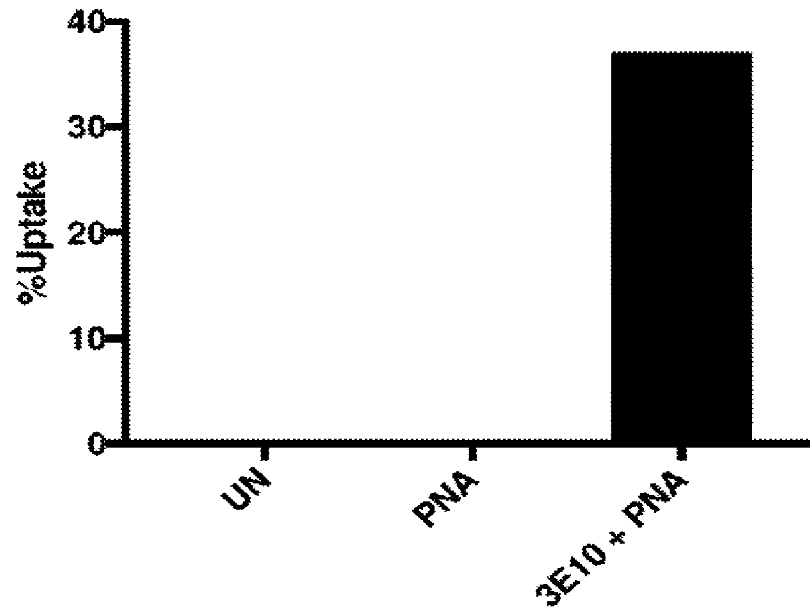


FIG. 2D

Untreated

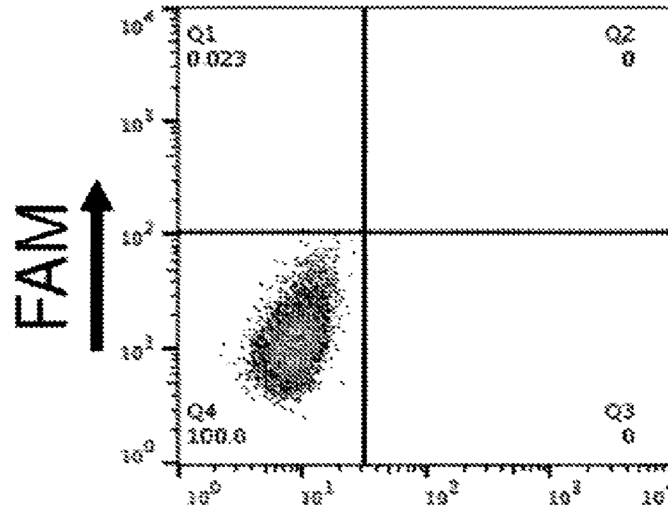


FIG. 3A

siRNA Only

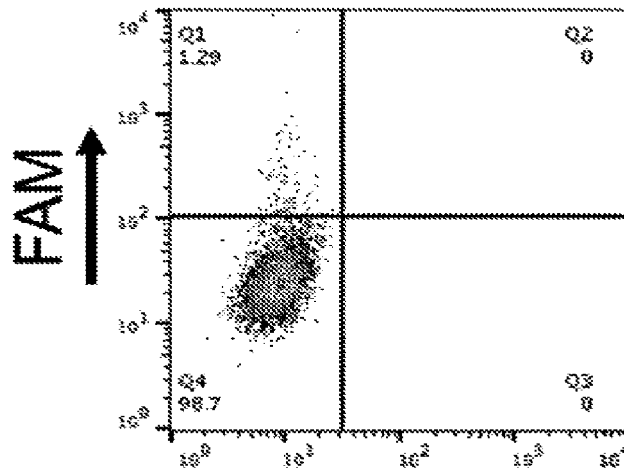


FIG. 3B

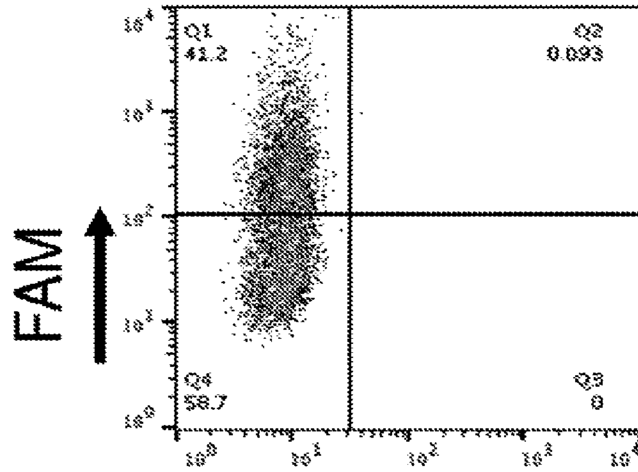


FIG. 3C

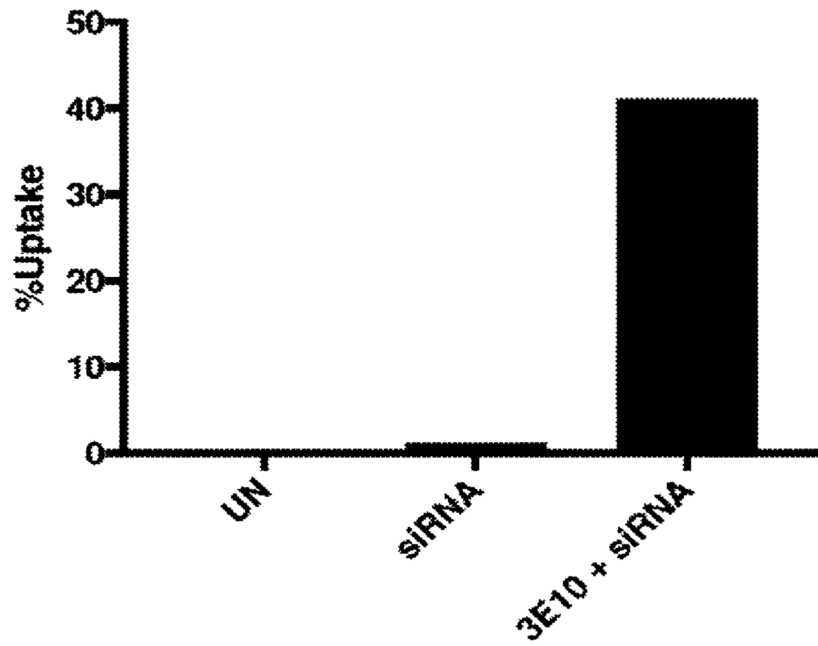


FIG. 3D

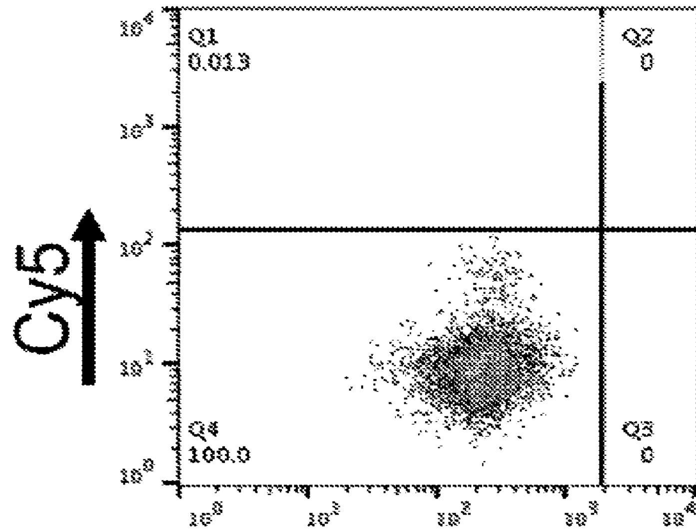


FIG. 4A

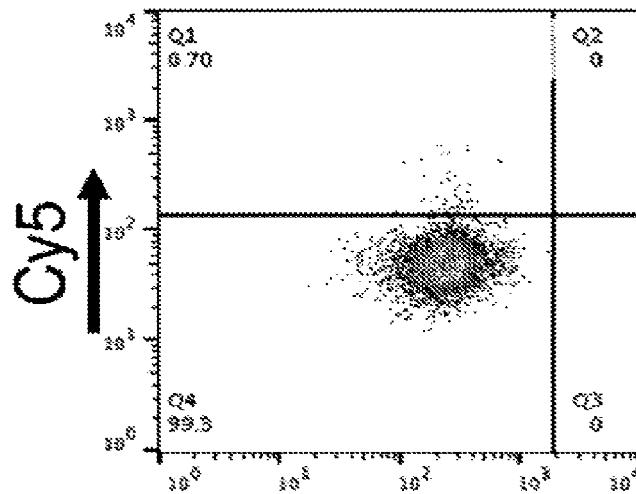


FIG. 4B

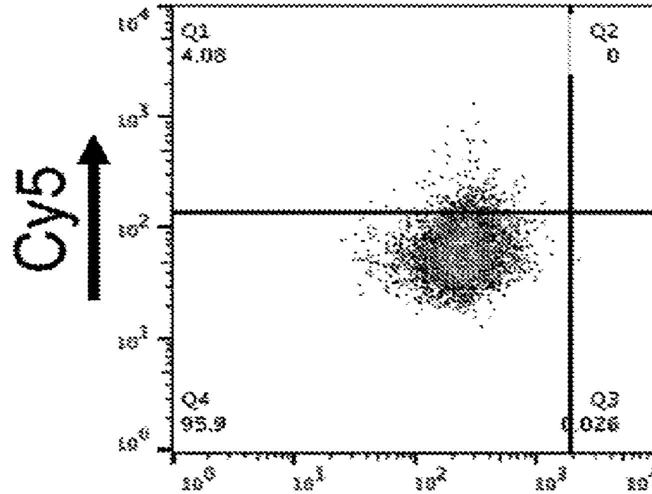


FIG. 4C

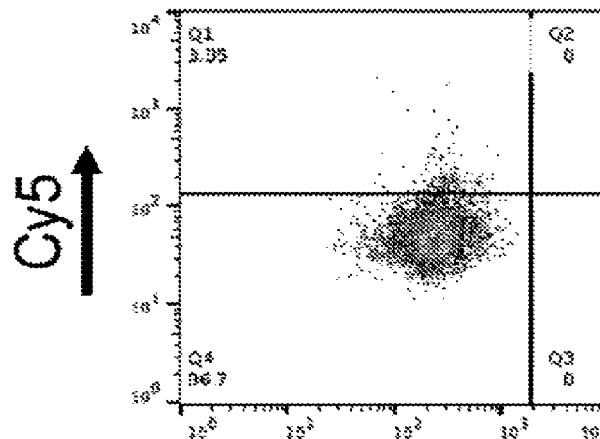


FIG. 4D

mRNA + 3E10 (10 uM)

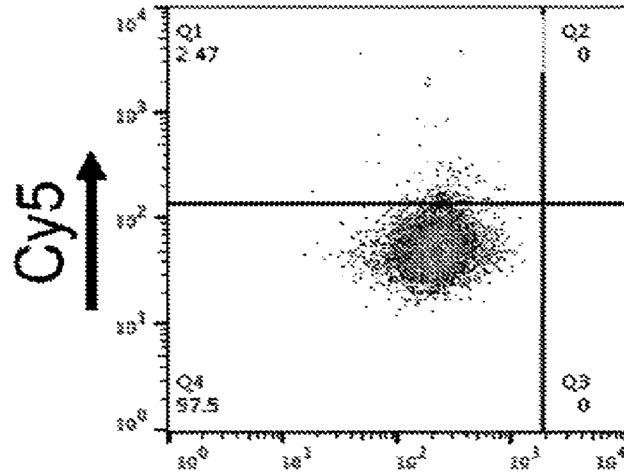


FIG. 4E

mRNA + 3E10 D31N (2.5 uM)

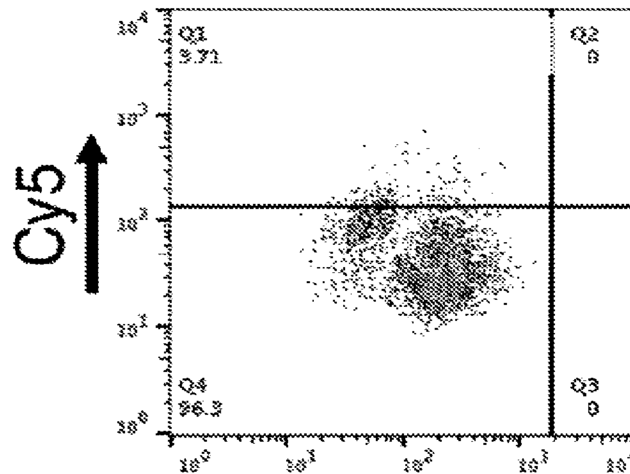


FIG. 4F

mRNA + 3E10 D31N (5 uM)

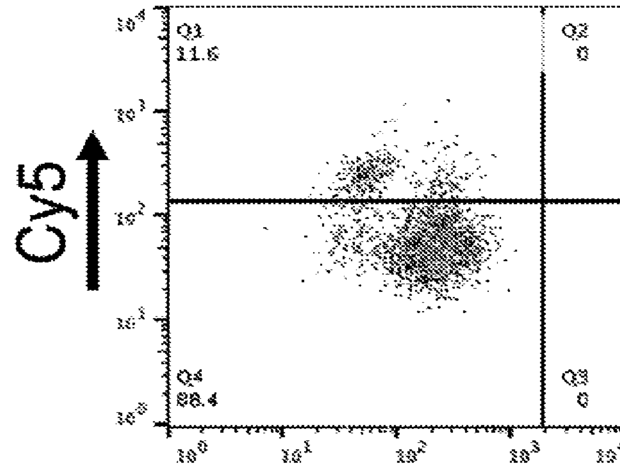


FIG. 4G

mRNA + 3E10 D31N (10 uM)

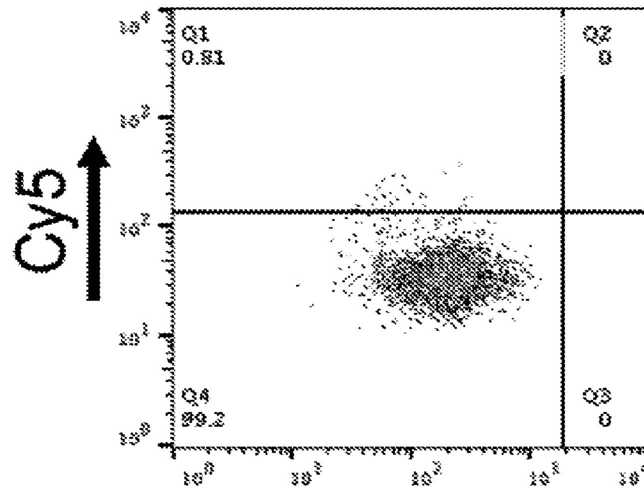


FIG. 4H

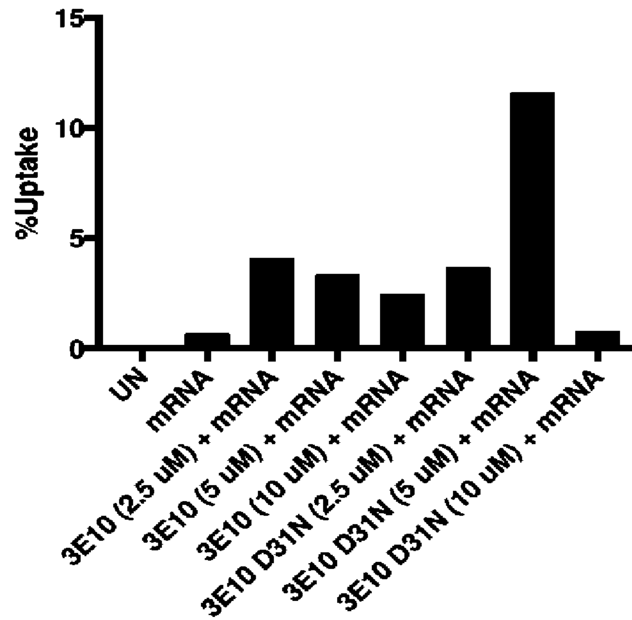


FIG. 4I

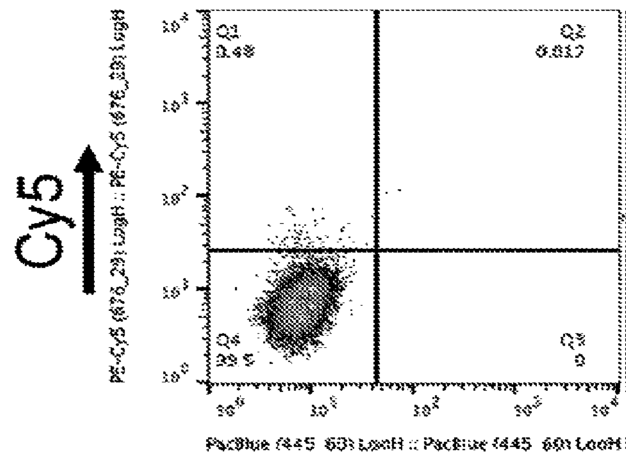


FIG. 5A

mRNA Only

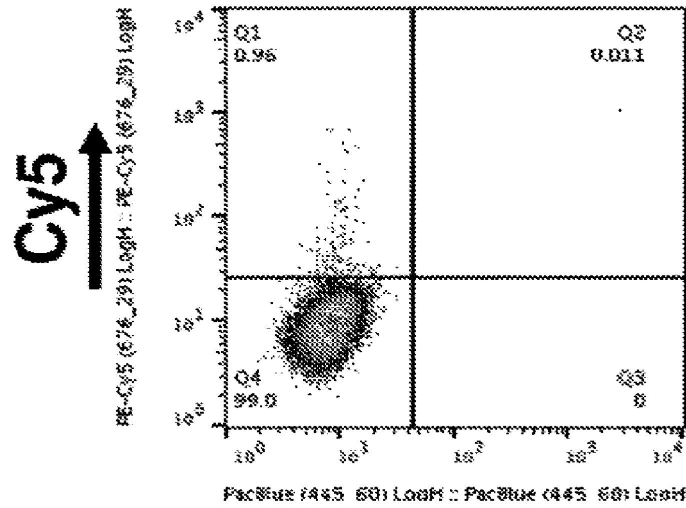


FIG. 5B

mRNA + 3E10 D31N (0.1 uM)

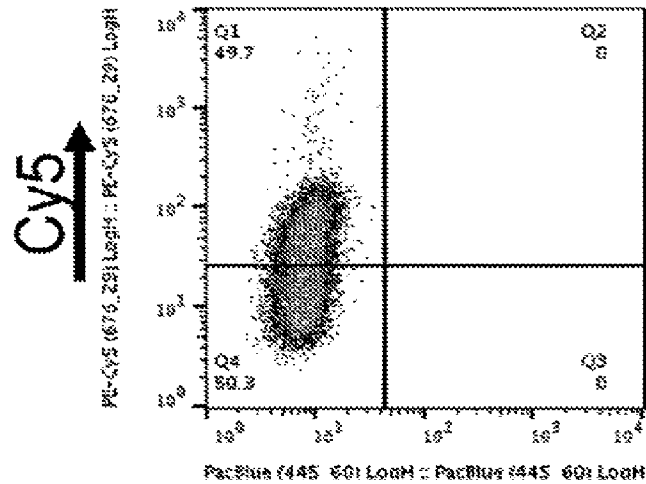


FIG. 5C

mRNA + 3E10 D31N (0.25 uM)

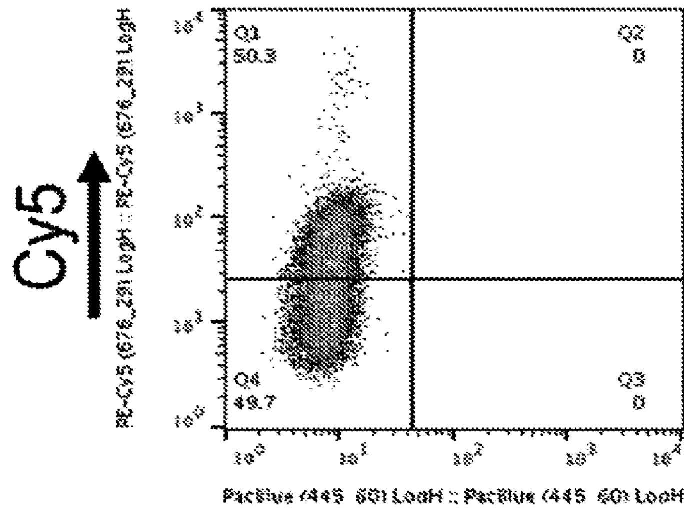


FIG. 5D

mRNA + 3E10 D31N (0.5 uM)

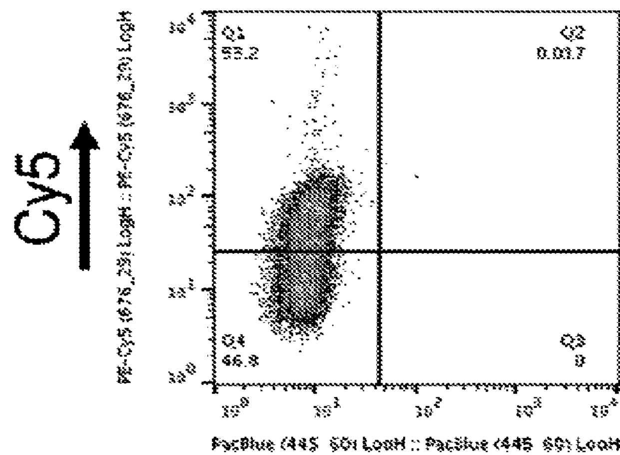


FIG. 5E

mRNA + 3E10 D31N (1 uM)

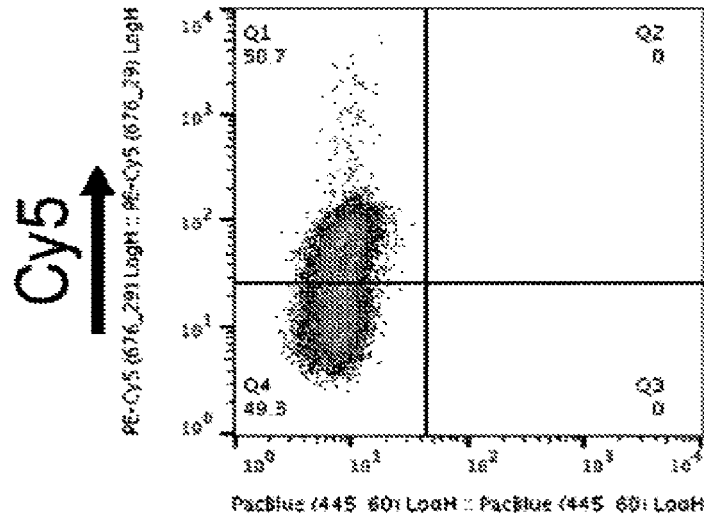


FIG. 5F

mRNA + 3E10 D31N (5 uM)

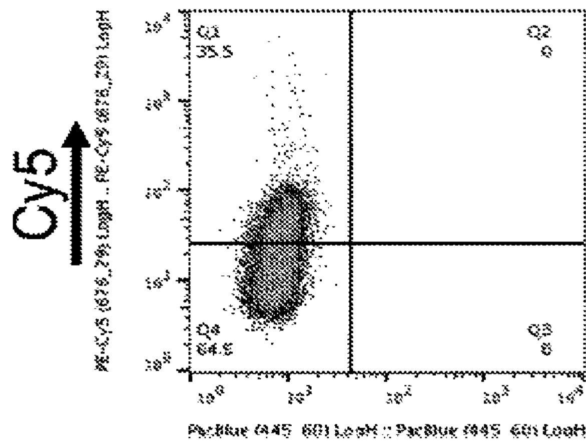


FIG. 5G

mRNA + 3E10 D31N (10 uM)

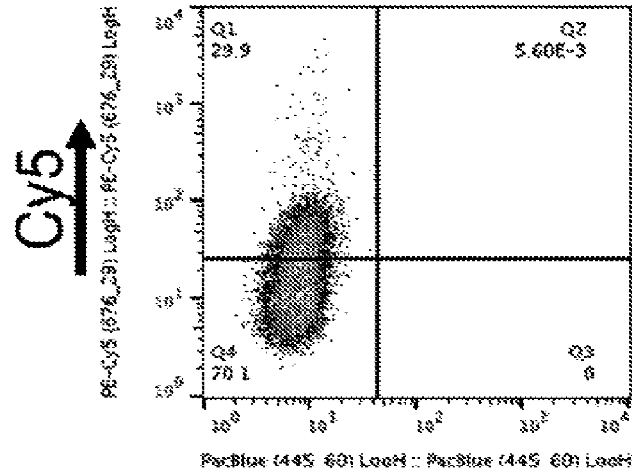


FIG. 5H

3E10 D31N mRNA Delivery

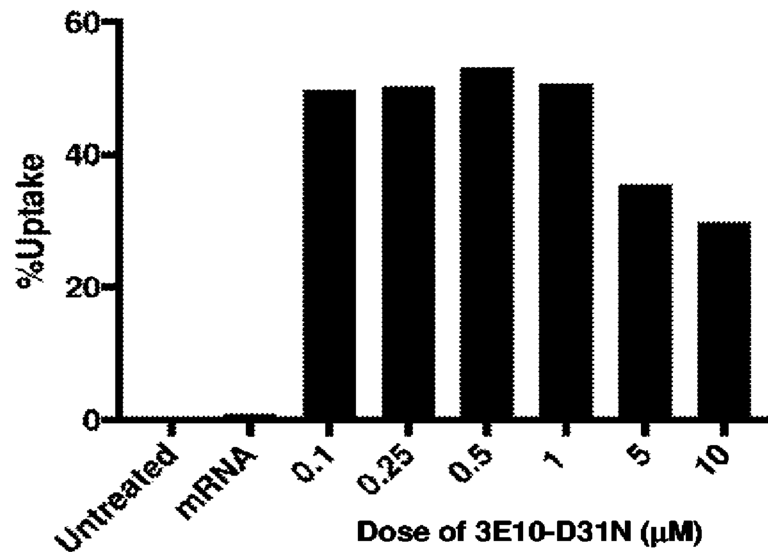


FIG. 5I

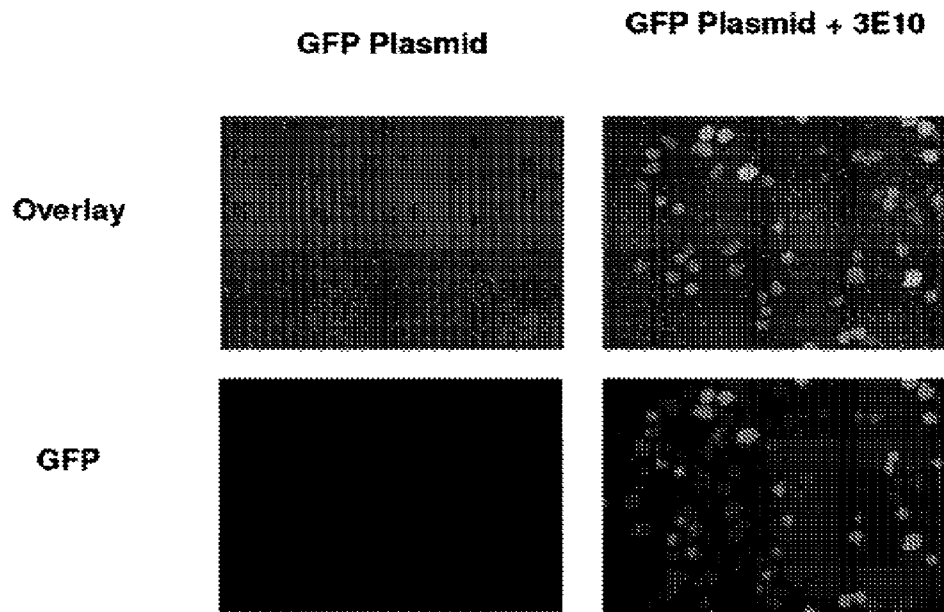


FIG. 6

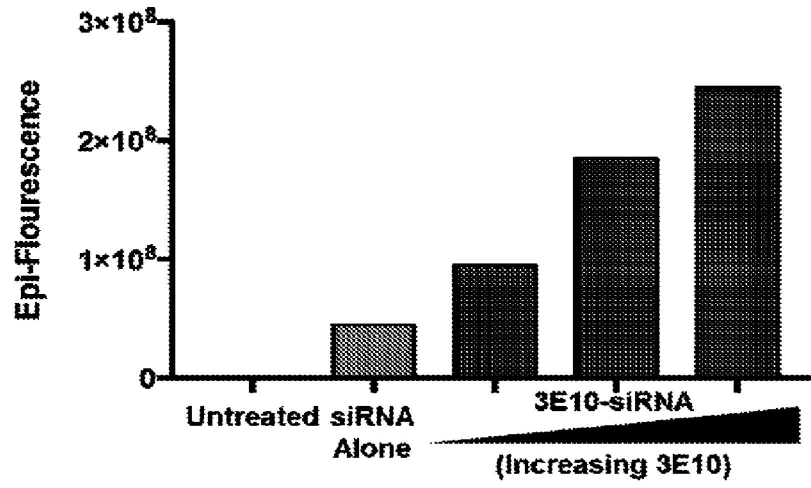


FIG. 7A

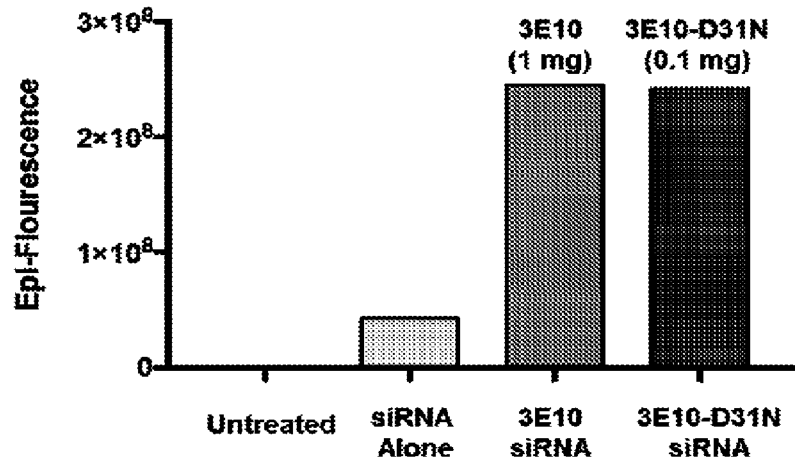


FIG. 7B

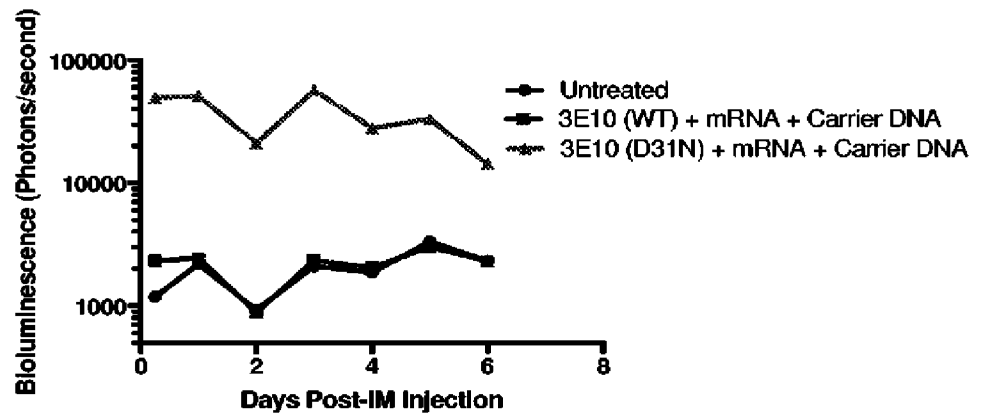
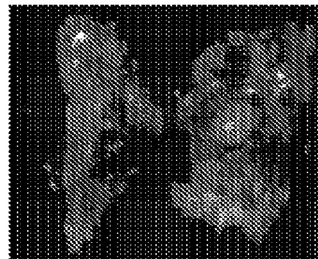


FIG. 8

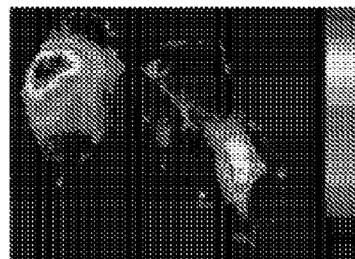
Untreated Control



Mouse muscle

FIG9A

3E10-D31N



Mouse muscle

FIG. 9B

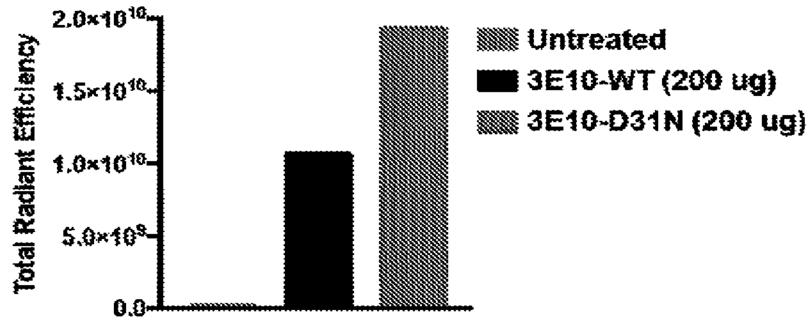


FIG. 9C

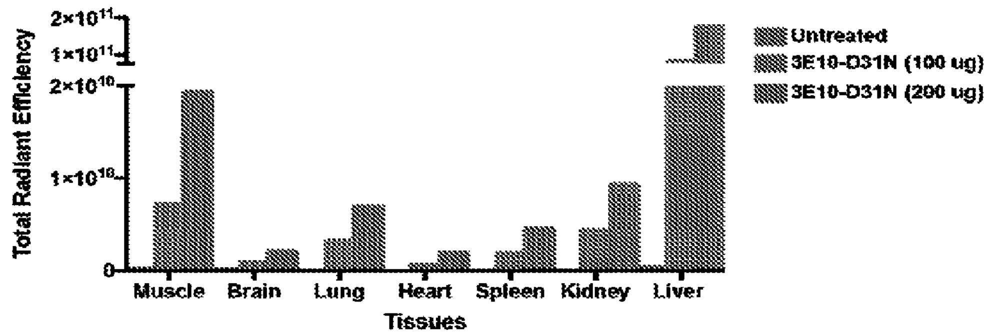


FIG. 10

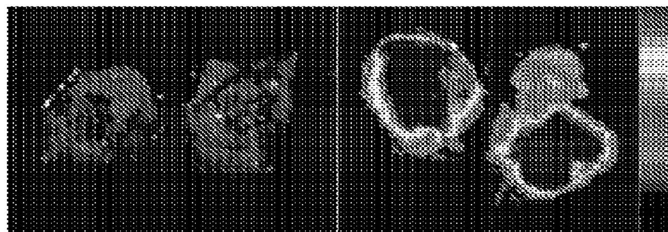


FIG. 11A FIG. 11B

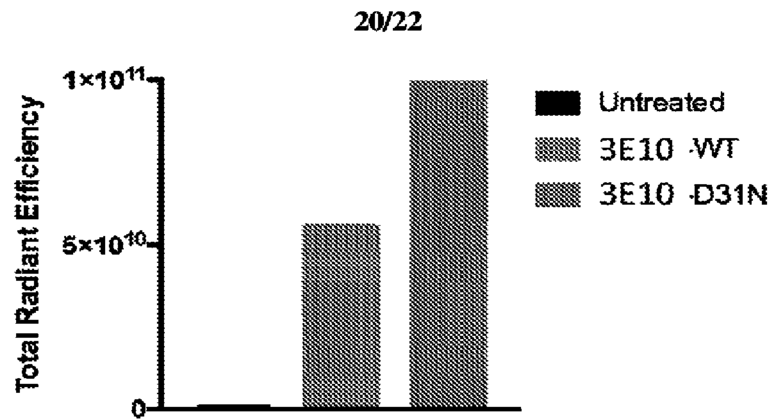


FIG. 11C

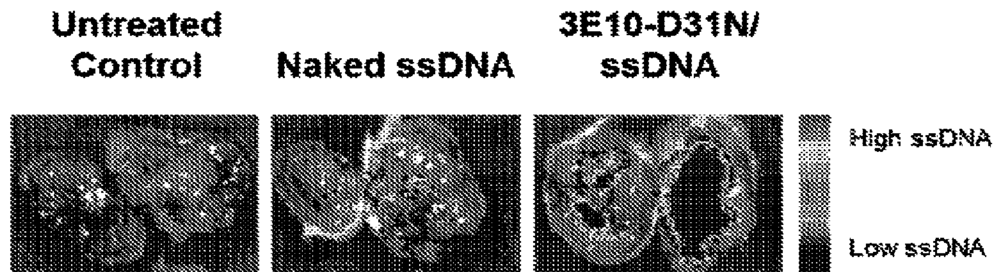


FIG. 12A FIG. 12B FIG. 12C

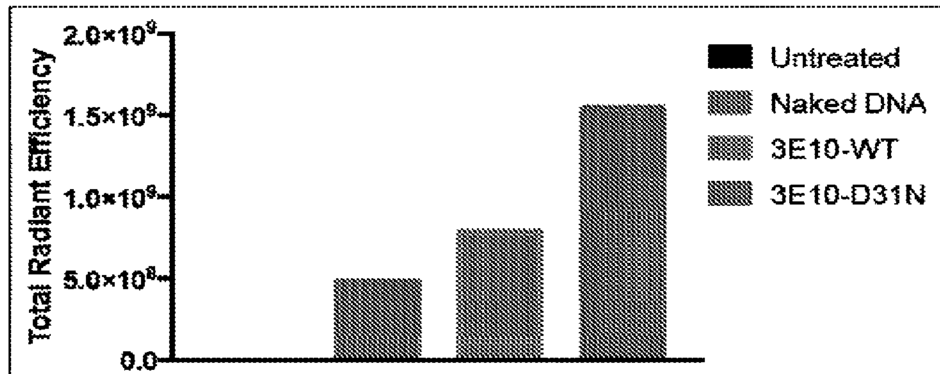


FIG. 12D

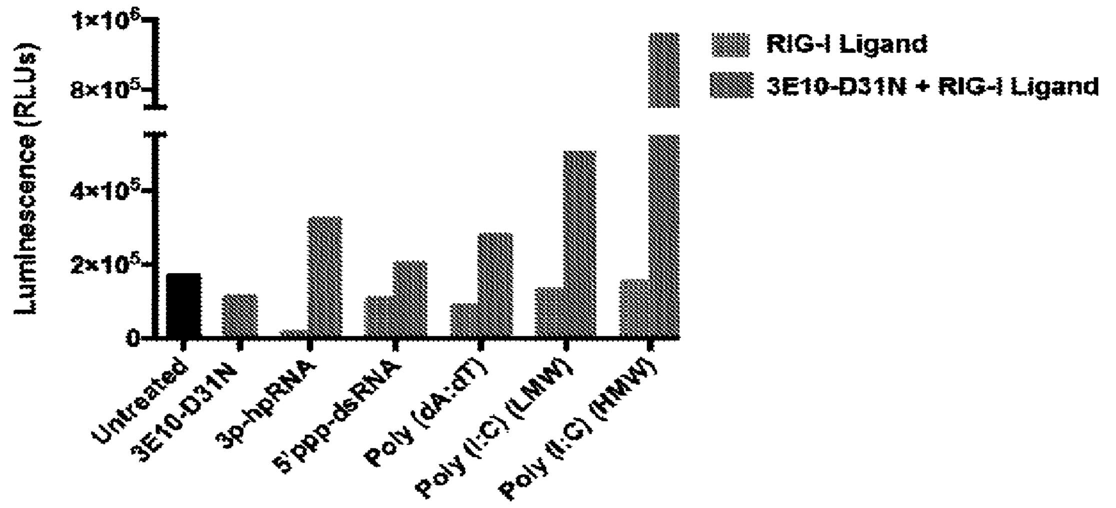


FIG. 13

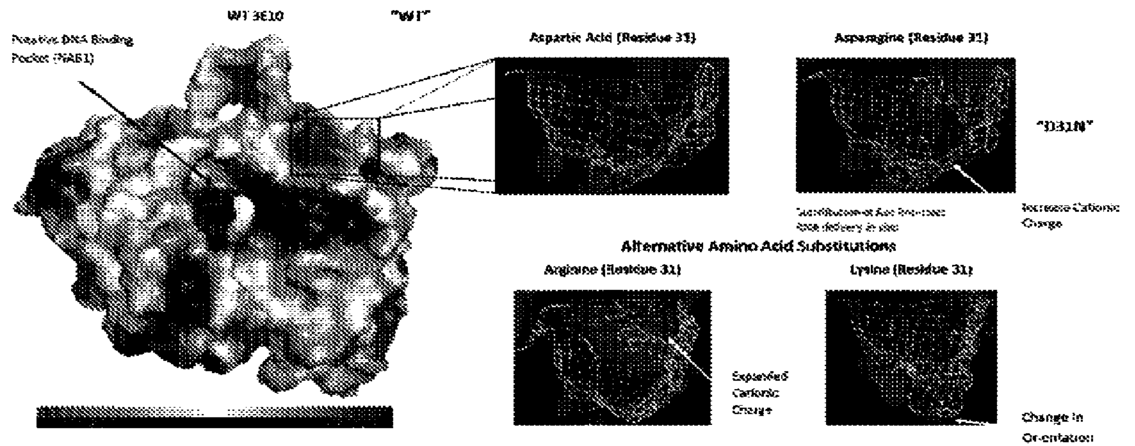


FIG. 14A

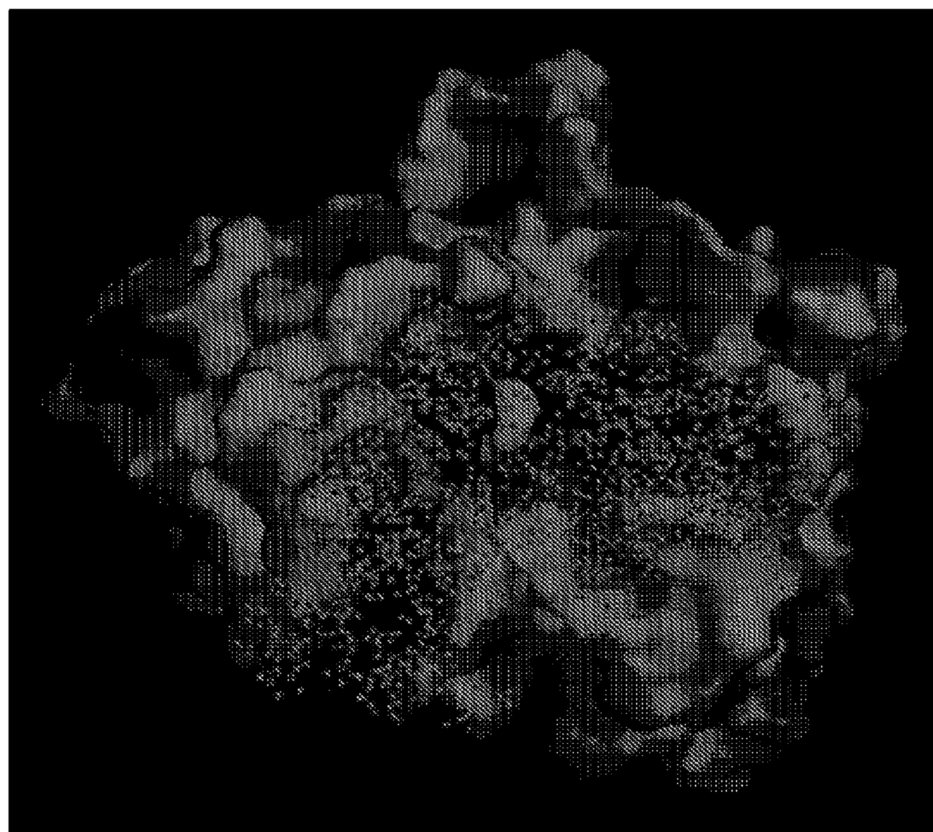


FIG. 14B