COMPOSITIONS AND METHODS FOR THE TREATMENT OF MYELOPROLIFERATIVE NEOPLASMS

CROSS-REFERENCE TO RELATED APPLICATION

[0001] This application claims priority to, and benefit of, U.S. Provisional Application No. 63/569,841, filed on March 26, 2024, the contents of which are incorporated by reference in their entirety herein.

REFERENCE TO AN ELECTRONIC SEQUENCE LISTING

[0002] The contents of the electronic sequence listing (REGE-030_001WO_SeqList_ST26.xml; Size: 163,736 bytes; and Date of Creation: March 13, 2025) are herein incorporated by reference in their entireties.

BACKGROUND

[0003] Mutations expressed by pre-malignant neoplasms and cancers represent a class of neoantigens for targeting by therapeutic vaccination, such as therapeutic DNA vaccination.

[0004] Targeting of protein antigens to antigen-presenting cells (APCs) can improve T-cell and B-cell responses. Recombinant immunoglobulin (Ig) molecules are particularly suited to this purpose. For example, short antigenic epitopes can replace loops between β-strands in the Ig constant domains, while targeted antigen delivery is obtained by equipping the recombinant Ig with variable (V) regions specific for surface molecules on APCs. However, such a strategy is problematic for larger antigens containing unidentified epitopes, and recombinant Ig molecules with short T cell epitopes fail to elicit antibodies against conformational epitopes. To overcome these limitations, targeted Ig-based homodimeric DNA vaccines known as Vaccibodies™ can be used that express infectious or tumor antigens.

[0005] Calreticulin, also called CALR, is a conserved chaperone protein that regulates numerous cellular processes, including protein folding, cell adhesion, and calcium homeostasis. Mutations in calreticulin, particularly frameshift mutations in exon 9, are common driver mutations in myeloproliferative neoplasms (MPN).

[0006] Therefore, there is a need for new and useful compositions, such as vaccinations, targeting calreticulin for use in treatment of cancer and related diseases.

SUMMARY

[0007] In one aspect, provided herein is a pharmaceutical composition comprising a polynucleotide comprising one or more sequences encoding a targeting unit, one or more sequences encoding a dimerization unit, and one or more sequences encoding an antigenic unit, wherein the antigenic unit comprises a sequence encoding at least one epitope derived from calreticulin, and a pharmaceutically acceptable carrier.

[0008] In some embodiments, the at least one epitope derived from calreticulin comprises a sequence of a frameshift mutation in calreticulin exon 9.

[0009] In some embodiments, the at least one epitope derived from calreticulin comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 4-11. In some embodiments, the frameshift mutation in calreticulin comprises an amino acid sequence of SEQ ID NO: 2-3 or 141-158. In some embodiments, the frameshift mutation in calreticulin comprises an amino acid sequence of SEQ ID NO: 3. In some embodiments, a polynucleotide encoding the frameshift mutation in calreticulin comprises a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical to SEQ ID NO:183. In some embodiments, a polynucleotide encoding the frameshift mutation in calreticulin comprises SEQ ID NO:183, or a sequence with 1, 2, 3, 4 or 5 insertions, substitutions or deletions relative thereto.

[0010] In some embodiments, the at least one epitope has a length of 7 to 44 amino acids. In some embodiments, the antigenic unit comprises at least two, at least three, at least four, or at least five epitopes from calreticulin.

[0011] In some embodiments, the dimerization unit comprises a hinge region. In some embodiments, the hinge region is derived from an immunoglobulin.

[0012] In some embodiments, the dimerization unit further comprises a dimerization domain. In some embodiments, the dimerization domain comprises an immunoglobulin constant domain. In some embodiments, the immunoglobulin constant domain is a carboxyterminal C domain derived from an IgG. In some embodiments, the carboxyterminal C domain is derived from IgG3.

[0013] In some embodiments, the dimerization unit comprises a linker. In some embodiments, the linker connects the hinge region and the dimerization domain. In some embodiments, the dimerization unit comprises hinge exons h1 and h4 connected through a linker to a CH3 domain of human lgG3.

[0014] In some embodiments, the dimerization unit comprises a sequence having at least 80%, at least 90%, at least 95%, or at least 99% identity to SEQ ID NO:178. In some embodiments, the dimerization unit comprises a sequence of SEQ ID NO:178. In some embodiments, a polynucleotide encoding the dimerization unit comprises a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical to SEQ ID NO:185. In some embodiments, the nucleotide sequence comprises a sequence of SEQ ID NO:185 or a sequence with 1, 2, 3, 4 or 5 insertions, substitutions or deletions relative thereto.

[0015] In some embodiments, the antigenic unit and the dimerization unit are connected by a linker. In some embodiments, the linker comprises an amino acid sequence of SEQ ID NO:98. In some embodiments, a polynucleotide encoding the linker comprises a sequence of SEQ ID NO:187.

[0016] In some embodiments, the targeting unit comprises a moiety that interacts with a surface molecule on an antigen presenting cell. In some embodiments, the surface molecule is selected from the group consisting of HLA, CD14, CD40, a chemokine receptor, and a Toll-like receptor. In some embodiments, the targeting unit comprises a MIP-1α chemokine that binds to a CCR5 receptor. In some embodiments, the MIP-1α chemokine comprises an amino acid sequence of SEQ ID NO:159. In some embodiments, a polynucleotide encoding the MIP-1α chemokine comprises a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical to SEQ ID NO:184. In some embodiments, the nucleotide sequence comprises a sequence of SEQ ID NO:184 or a sequence with 1, 2, 3, 4 or 5 insertions, substitutions or deletions relative thereto.

[0017] In some embodiments, the polynucleotide is an RNA or a DNA.

[0018] In some embodiments, the polynucleotide comprises a nucleotide sequence encoding a signal peptide. In some embodiments, the signal peptide comprises SEQ ID NO:163. In some embodiments, a polynucleotide encoding the signal peptide comprises a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical to SEQ ID NO:186. In some embodiments, the nucleotide sequence comprises a sequence of SEQ ID NO:186 or a sequence with 1, 2, 3, 4 or 5 insertions, substitutions or deletions relative thereto.

[0019] In some embodiments, the polynucleotide comprises, from 5′ to 3′, one or more sequences encoding the targeting unit, the dimerization unit, and the antigenic unit. In some embodiments, the targeting unit encodes a MIP-1α sequence, the dimerization unit encodes a

hinge exon h1 from IgG3, a hinge exon h4 from IgG3, a dimerization unit linker, and human CH3 domain of IgG3, and the antigenic unit encodes an epitope derived from calreticulin comprising SEQ ID NOs:4-11. In some embodiments, the polynucleotide comprises a sequence encoding a signal peptide of MIP-1α 5′ to the targeting unit, and a sequence encoding a linker between the dimerization unit and the antigenic unit.

[0020] In some embodiments, the polynucleotide comprises, from 5′ to 3′: a targeting unit encoding SEQ ID NO:159; a dimerization unit encoding SEQ ID NO:178; and an antigenic unit encoding SEQ ID NO:3. In some embodiments, the polynucleotide comprises, from 5′ to 3′: a signal peptide encoding SEQ ID NO:163; a targeting unit encoding SEQ ID NO:159; a dimerization unit encoding SEQ ID NO:178; a linker encoding SEQ ID NO:98; and an antigenic unit encoding SEQ ID NO:3.

[0021] In some embodiments, the polynucleotide comprises, from 5′ to 3′: SEQ ID NO:184; SEQ ID NO:185; and SEQ ID NO: 183. In some embodiments, the polynucleotide comprises, from 5′ to 3′: SEQ ID NO:186; SEQ ID NO:184; SEQ ID NO:185; SEQ ID NO:187; and SEQ ID NO: 183.

[0022] In another aspect, provided herein is a pharmaceutical composition comprising a polynucleotide encoding an amino acid sequence of SEQ ID NO:177.

[0023] In another aspect, provided herein is a polynucleotide of any one of the above aspects or embodiments.

[0024] In another aspect, provided herein is an expression vector comprising the polynucleotide of the above aspect.

[0025] In another aspect, provided herein is a host cell comprising the polynucleotide or the expression vector of any of the above aspects.

[0026] In another aspect, provided herein is a polypeptide encoded by the polynucleotide of any of the above aspects or embodiments.

[0027] In another aspect, provided herein is a dimeric protein comprising two polypeptides of the above aspect. In some embodiments, the dimeric protein is a homodimer.

[0028] In some embodiments, the polynucleotide, the polypeptide, or the dimeric protein of the above aspects or embodiments is for use in the manufacture of a medicament for treating a cancer. In some embodiments, the cancer is a myeloproliferative neoplasm.

[0029] In some embodiments, the pharmaceutical composition of any one of the above aspects or embodiments is for use in the manufacture of a medicament for treating cancer in a subject. In some embodiments, the cancer is a myeloproliferative neoplasm.

[0030] In another aspect, provided herein is a method of treating cancer in a subject in need thereof, comprising administering a therapeutically effective amount of the pharmaceutical composition of any one of the above aspects or embodiments to the subject. In some embodiments, the cancer is a myeloproliferative neoplasm.

[0031] In another aspect, provided herein is a method of eliciting an immune response in a subject in need thereof, comprising administering the pharmaceutical composition of any one of the above aspects or embodiments to the subject.

[0032] In another aspect, provided herein is a method of treating cancer in a subject in need thereof, the method comprising administering to the subject a pharmaceutical composition comprising a polynucleotide comprising one or more sequences encoding a targeting unit, one or more sequences encoding a dimerization unit, and one or more sequences encoding an antigenic unit, wherein the antigenic unit comprises a sequence encoding at least one epitope derived from calreticulin, and a pharmaceutically acceptable carrier. In some embodiments, the cancer is a myeloproliferative neoplasm.

[0033] In some embodiments, the myeloproliferative neoplasm is chronic myeloid leukemia, acute myeloid leukemia, polycythaemia vera, essential thrombocythemia, primary myelofibrosis, chronic eosinophilic leukemia, chronic myelomonocytic leukemia, systemic mastocytosis, idiopathic myelofibrosis, or myeloma.

[0034] In some embodiments, the cancer expresses calreticulin. In some embodiments, the calreticulin comprises a frameshift mutation in exon 9.

[0035] In some embodiments, the treatment is therapeutic.

[0036] In some embodiments, the treatment is prophylactic.

DESCRIPTION OF THE DRAWINGS

[0037] The features of the present disclosure are set forth with particularity in the appended claims. A better understanding of the features and advantages of the present disclosure will be obtained by reference to the following detailed description that sets forth illustrative embodiments, in which the principles of the disclosure are utilized, and the accompanying drawings of which:

[0038] **FIG. 1A** shows the results of immunopeptidomic mass spectrometry (IP-MS) for 9-mers of calreticulin and distribution of peptide length. Peptide count is indicated on the y-axis, which indicates from bottom to top: 0, 1000, 2000, 3000, 4000. Peptide length (from 7 to 20 amino acids) is indicated on the x-axis.

[0039] **FIG. 1B** depicts mass spectra of the endogenous or stable isotope labeled (heavy Spike-in) peptides. SPARPRTSC: SEQ ID NO:5.

[0040] **FIG. 2A** is a schematic depiction of an example protocol for antigen-specific memory T cell expansion. PBMC: peripheral blood mononuclear cells; IFNα: interferon-alpha; GM-CSF: Granulocyte-macrophage colony-stimulating factor; IL: interleukin.

[0041] **FIG. 2B** depicts flow cytometry plots for myeloproliferative neoplasm (MPN) patient and healthy donor $CD8^+$ T cells upon stimulation with DMSO, a wild type calreticulin ($CALR^{WT}$) peptide pool, a mutant calreticulin comprising the exon 9 frameshift ($CALR^{MUT}$) peptide pool, or CEF viral peptide pool (Cytomegalovirus, Epstein-Barr virus, and Influenza, abbreviated CMV, EBV, and Flu respectively).

[0042] **FIG. 2C** is a graph depicting a summary of IFNγ and TNFα production from MPN patient (right plot) or healthy donor (left plot) CD8+ T cells upon stimulation with DMSO, a $CALR^{MUT}$ peptide pool, a $CALR^{WT}$ peptide pool, or a CEF pool.

[0043] **FIG. 3A** is a schematic depicting an example protocol for antigen-specific naïve T cell expansion. TIIP: TNFα, IL-1, IL-6, PGE1 (Prostaglandin E1); moDC: monocyte-derived dendritic cell.

[0044] **FIG. 3B** is a series of immunofluorescence microscopy images depicting IFNγ production using an IFNγ ELISpot assay. Naïve T cells from Healthy Donor 13 (HD13) were expanded with DMSO or $CALR^{MUT}$ peptides for 26 days, then stimulated with the indicated peptide pools or DMSO.

[0045] **FIG. 3C** is a graph depicting IFNγ production measured using an IFNγ ELISpot assay. Naïve T cells from four donors were expanded with DMSO or $CALR^{MUT}$ peptides for 20 days, then stimulated with a $CALR^{MUT}$ peptide pool, $CALR^{WT}$ peptide pool, or DMSO. HD: human donor.

[0046] **FIG. 4** provides box plots depicting IFNγ ELISpot assay results for CALR44mer-specific T cells from splenocytes of C57BL/6 mice stimulated overnight with a 15-mer peptide pool covering the entire CALR44mer. Animals were dosed with either the CALR44mer Vaccibody™ (at 20 μg or 60 μg doses), or empty Vaccibody™ (negative control). The y-axis depicts IFNγ spot forming units (SFU) per 10^6 splenocytes. Individual symbols correspond to individual mice (n=5 mice per group), and boxes and whiskers depict median, lower and upper quartiles, and lowest and highest values.

[0047] **FIG. 5A** depicts a schematic of the construct used to engineer TC-1 cells expressing a HA-Ub-CALR^{MUT}44-mer construct (left), and flow cytometry validation of intracellular expression in engineered TC-1 cells by HA-tag (HA) detection (right).

[0048] **FIG. 5B** provides line graphs depicting *in vivo* tumor control of the HA-Ub-CALR-44mer engineered TC1 cell line in C57BL/6 mice, vaccinated with Empty Vaccibody™ (negative control) or CALR^{MUT} 44-mer Vaccibody™, with mice immunized twice by intramuscular injection followed by electroporation at days 7 and 14 post cell implantation. Tumor growth curves are shown for individual animals (n=5 per group). Tumor volume (in mm^2) is depicted on the y-axis, and days after implantation is depicted on the x-axis.

DETAILED DESCRIPTION

[0049] The present disclosure provides a pharmaceutical composition comprising a polynucleotide comprising one or more sequences encoding a targeting unit, one or more sequences encoding a dimerization unit, and one or more sequences encoding an antigenic unit derived from calreticulin. Such polynucleotides, *e.g.*, a Vaccibody™ construct, are directed to neoantigens derived from calreticulin, with uses for prophylaxis, to reduce the likelihood of, or treatment of malignant neoplasms or cancer. As described, for example, in WO 2022/238363, incorporated herein by reference in its entirety, a Vaccibody™ is a multimeric protein containing multiple polypeptides, such as a dimeric protein comprising or consisting of two polypeptides, each a fusion protein comprising a targeting unit, which targets or is capable of targeting antigen-presenting cells, a multimerization unit, such as a dimerization unit, and an antigenic unit.

[0050] Compositions according to the present disclosure are designed for evoking a cell-mediated immune response, *e.g.*, through activation of T cells against the neoantigens as described herein.

[0051] Unless defined otherwise, technical and scientific terms used herein have meanings that are commonly understood by those of ordinary skill in the art unless defined otherwise. Generally, terminologies pertaining to techniques of molecular biology, nucleic acid chemistry, protein chemistry, genetics, microbiology, transgenic cell production, and hybridization described herein are those well-known and commonly used in the art. Techniques and procedures described herein are generally performed according to conventional methods well known in the art and as described in various general and more specific references that are cited and discussed throughout the instant specification. For example, *see* Sambrook et al., Molecular Cloning: A Laboratory Manual (Third ed., Cold

Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. 2000). *See also* Ausubel *et al.*, Current Protocols in Molecular Biology, Greene Publishing Associates (1992). The nomenclatures utilized in connection with, and the laboratory procedures and techniques described herein are those well-known and commonly used in the art unless otherwise specified.

[0052] Unless otherwise required by context herein, singular terms shall include pluralities and plural terms shall include the singular. Singular forms "a", "an" and "the", and singular use of any word, include plural referents unless expressly and unequivocally limited on one referent.

[0053] It is understood the use of the alternative term (*e.g.*, "or") is taken to mean either one or both or any combination thereof of the alternatives.

[0054] The term "and/or" used herein is to be taken to mean specific disclosure of each of the specified features or components with or without the other. For example, the term "and/or" as used in a phrase such as "A and/or B" herein is intended to include: "A and B"; "A or B"; "A" (A alone); and "B" (B alone). In a similar manner, the term "and/or" as used in a phrase such as "A, B, and/or C" is intended to encompass each of the following aspects: "A, B, and C"; "A, B, or C"; "A or C"; "A or B"; "B or C"; "A and B"; "B and C"; "A and C"; "A" (A alone); "B" (B alone); and "C" (C alone).

[0055] As used herein and in the appended claims, the terms "comprising", "including", "having" and "containing", and their grammatical variants, as used herein, are intended to be non-limiting so that one item or multiple items in a list do not exclude other items that can be substituted or added to the listed items. It is understood that wherever aspects are described herein with the language "comprising," otherwise analogous aspects described in terms of "consisting of" and/or "consisting essentially of" are also provided.

[0056] As used herein, the terms "about" and "approximately" refer to a value or composition that is within an acceptable error range for the particular value or composition as determined by one of ordinary skill in the art, which will depend in part on how the value or composition is measured or determined, *i.e.*, the limitations of the measurement system. For example, "about" or "approximately" can mean within one or more than one standard deviation per the practice in the art. Alternatively, "about" or "approximately" can mean a range of up to 10% (*i.e.*, ±10%) or more depending on the limitations of the measurement system. For example, about 5 mg can include any number between 4.5 mg and 5.5 mg. Furthermore, particularly with respect to biological systems or processes, the terms can mean

up to an order of magnitude or up to 5-fold of a value. When particular values or compositions are provided in the instant disclosure, unless otherwise stated, the meaning of "about" or "approximately" should be assumed to be within an acceptable error range for that particular value or composition. Also, where ranges and/or subranges of values are provided, the ranges and/or subranges can include the endpoints of the ranges and/or subranges.

[0057] As used herein, calreticulin (also known as CALR, CALR1, calregulin, SSA, CRT, CRTC, RO, Autoantigen RO, FLJ26680, CC1qR, CRP55, ERp60, HACBP, Grp60, Epididymis Secretory Sperm Binding Protein Li 99n, and HEL-S-99n) refers to the protein of Uniprot Accession No. P27797, NCBI accession number NG_029662.1 (gene), and NP_004334.1 (protein), and related isoforms and orthologs. As used herein, "wildtype calreticulin" refers to a calreticulin polypeptide or protein comprising SEQ ID NO:1, or a polynucleotide encoding the same. As used herein, "frameshift calreticulin" or "calreticulin frameshift mutation" refers to a mutation in the wildtype calreticulin gene (*e.g.*, SEQ ID NO:1) leading to an alternative open reading frame (ORF). In some embodiments, a frameshift mutation in calreticulin occurs in exon 9 (SEQ ID NO:171). A frameshift mutation may be caused by deletion of one or more nucleotides or by insertion of one or more nucleotides.

[0058] As used herein, the terms "subject" and "patient" refer to an organism to be treated by the methods and compositions described herein. Such organisms preferably include, but are not limited to, mammals (*e.g.*, murines, simians, equines, bovines, porcines, canines, felines, and the like), and more preferably include humans.

[0059] The terms "nucleic acid", "polynucleotide" and "oligonucleotide" and other related terms used herein are used interchangeably and refer to polymers of nucleotides and are not limited to any particular length. Nucleic acids include recombinant and chemically-synthesized forms. Nucleic acids can be isolated. Nucleic acids include DNA molecules (*e.g.*, cDNA or genomic DNA), RNA molecules (*e.g.*, mRNA), analogs of the DNA or RNA generated using nucleotide analogs (*e.g.*, peptide nucleic acids (PNA) and non-naturally occurring nucleotide analogs), and chimeric forms containing DNA and RNA. Nucleic acids can be single-stranded or double-stranded. Nucleic acids comprise polymers of nucleotides, where the nucleotides may include natural or non-natural bases, and/or sugars. Nucleic acids comprise naturally-occurring internucleosidic linkages, for example phosphodiester linkages. Nucleic acids can lack a phosphate group. Nucleic acids can comprise non-natural internucleoside linkages, including phosphorothioate, phosphorothiolate, and/or peptide

nucleic acid (PNA) linkages. In some embodiments, nucleic acids comprise one type of polynucleotides or a mixture of two or more different types of polynucleotides.

[0060] As used herein, the term "effective amount" or "therapeutically effective amount" refers to the amount of a compound (*e.g.*, a nucleic acid or protein of the present disclosure) sufficient to effect beneficial or desired results. An effective amount can be administered in one or more administrations, applications or dosages and is not intended to be limited to a particular administration route.

[0061] As used herein, the term "treating" includes any effect, *e.g.*, lessening, reducing, modulating, ameliorating or eliminating, that results in the improvement of the condition, disease, disorder, and the like, or ameliorating a symptom thereof. In some embodiments, treatment may be preventative (*i.e.*, prophylactic, to reduce the likelihood of the condition, disease, disorder, and the like).

[0062] As used herein, the term "pharmaceutical composition" refers to the combination of an active agent with a carrier, inert or active, making the composition especially suitable for diagnostic or therapeutic use *in vivo* or *ex vivo*.

[0063] As used herein, the term "pharmaceutically acceptable carrier" refers to any of the standard pharmaceutical carriers, such as a phosphate buffered saline solution, water, emulsions (*e.g.*, such as an oil/water or water/oil emulsions), and various types of wetting agents. The compositions also can include stabilizers and preservatives. For examples of carriers, stabilizers and adjuvants, *see e.g.*, Martin, Remington's Pharmaceutical Sciences, 15th Ed., Mack Publ. Co., Easton, PA [1975].

[0064] As used herein, the term "pharmaceutically acceptable salt", refers to salts of such compounds that are appropriate for use in pharmaceutical contexts, *i.e.*, salts which are, within the scope of sound medical judgment, suitable for use in contact with the tissues of humans and lower animals without undue toxicity, irritation, allergic response and the like, and are commensurate with a reasonable benefit/risk ratio. Pharmaceutically acceptable salts are well known in the art. For example, S. M. Berge, *et al.* describes pharmaceutically acceptable salts in detail in J. Pharmaceutical Sciences, 66: 1-19 (1977).

[0065] As a general matter, compositions specifying a percentage are by weight unless otherwise specified. Further, if a variable is not accompanied by a definition, then the previous definition of the variable controls.

[0066] An antigenic unit of the polynucleotides of the present disclosure comprises one or more neoepitopes. Such neoepitopes, *e.g.*, an epitope derived from calreticulin, are

designed to be presented in MHC-neoepitope complexes. There are two primary classes of major histocompatibility complex (MHC) molecules, MHC I and MHC II.

[0067] MHC I is found on the cell surface of all nucleated cells in the body. One function of MHC I is to display peptides of non-self proteins from within cells to cytotoxic T cells. The MHC I complex-peptide complex is inserted into the plasma membrane of the cell presenting the peptide to the cytotoxic T cells, triggering activation of cytotoxic T cells against the particular MHC-peptide complex. The peptide is positioned in a groove in the MHC I molecule, allowing the peptide to be about 8-10 amino acids long.

[0068] MHC class II molecules are a family of molecules normally found only on antigen-presenting cells such as dendritic cells, mononuclear phagocytes, some endothelial cells, thymic epithelial cells, and B cells. As opposed to MHC I, the antigens presented by MHC class II peptides are derived from extracellular proteins. Extracellular proteins are endocytosed, digested in lysosomes, and the resulting antigenic peptides are loaded onto MHC class II molecules and then presented at the cell surface. The antigen-binding groove of MHC class II molecules is open at both ends and is able to present longer peptides than MHC I, generally between 15 and 24 amino acid residues long.

[0069] Class I MHC molecules are recognized by CD8 and co-receptors on T cells, normally called CD8+ T cells (or CD8+ cells), whereas class II MHC molecules are recognized by CD4 and co-receptors on the T cells, normally called CD4+ T cells (or CD4+ cells).

[0070] Sequence identity may be determined by any suitable methods known in the art. An example method for determining sequence identity is as follows: a high level of sequence identity indicates likelihood that the first sequence is derived from the second sequence. Amino acid sequence identity requires identical amino acid sequences between two aligned sequences. Thus, a candidate sequence sharing 70% amino acid identity with a reference sequence requires that, following alignment, 70% of the amino acids in the candidate sequence are identical to the corresponding amino acids in the reference sequence. Identity may be determined by aid of computer analysis, such as, without limitations, the ClustalW computer alignment program (Higgins D., Thompson J., Gibson T., Thompson J.D., Higgins D.G., Gibson T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22:4673-4680), and the default parameters suggested therein. Using this program with its default settings, the mature (bioactive) part of a query

and a reference polypeptide are aligned. The number of fully conserved residues is counted and divided by the length of the reference polypeptide. In doing so, any tags or fusion protein sequences, which form part of the query sequence, are disregarded in the alignment and subsequent determination of sequence identity.

[0071] The ClustalW algorithm may similarly be used to align nucleotide sequences. Sequence identities may be calculated in a similar way as indicated for amino acid sequences.

[0072] Another non-limiting example of a mathematical algorithm utilized for the comparison of sequences is the algorithm of Myers and Miller, CABIOS (1989). Such an algorithm is incorporated into the ALIGN program (version 2.0) which is part of the FASTA sequence alignment software package (Pearson WR, Methods Mol Biol, 2000, 132: 185-219). Align calculates sequence identities based on a global alignment. AlignO does not penalize the gaps in the end of the sequences. When utilizing the ALIGN or AlignO program for comparing amino acid sequences, a BLOSUM50 substitution matrix with gap opening/extension penalties of -12/- 2 may be used.

[0073] Amino acid sequence variants of the amino acid sequences disclosed herein are contemplated as within the scope of the instant disclosure. Amino acid sequence variants may be prepared by introducing appropriate changes into the nucleotide sequence encoding the composition, or by peptide synthesis. Such modifications include, for example, deletions from, and/or insertions into and/or substitutions of, residues within the amino acid sequences. The terms substituted/substitution, deleted/deletions and inserted/insertions as used herein in reference to amino acid sequences and sequence identities are well known and clear to the skilled person in the art. Any combination of deletion, insertion, and substitution can be made to arrive at the final construct, provided that the final construct possesses the desired characteristics. For example, deletions, insertions, or substitutions of amino acid residues may produce a silent change and result in a functionally equivalent peptide/polypeptide.

[0074] Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues as long as the secondary binding activity of the substance is retained. For example, negatively charged amino acids include aspartic acid and glutamic acid; positively charged amino acids include lysine and arginine; and amino acids with uncharged polar head groups having similar hydrophilicity values include leucine, isoleucine, valine, glycine, alanine, asparagine, glutamine, serine, threonine, phenylalanine, and tyrosine. Herein encompassed are conservative substitutions, *i.e.*, like-for-like substitution such as basic for

basic, acidic for acidic, polar for polar *etc*. and non-conservative substitutions, *i.e*., from one class of residue to another or alternatively involving the inclusion of unnatural amino acids such as ornithine, diaminobutyric acid ornithine, norleucine, ornithine, pyriylalanine, thienylalanine, naphthyl alanine and phenylglycine. Conservative substitutions that may be made are, for example within the groups of basic amino acids (*e.g*., arginine, lysine and histidine), acidic amino acids (*e.g*., glutamic acid and aspartic acid), aliphatic amino acids (*e.g*., alanine, aaline, leucine, isoleucine), polar amino acids (*e.g*., glutamine, asparagine, serine, threonine), aromatic amino acids (*e.g*., phenylalanine, tryptophan, tyrosine), hydroxyl amino acids (*e.g*., serine, threonine), large amino acids (*e.g*., phenylalanine, tryptophan) and small amino acids (*e.g*., glycine, alanine).

[0075] Throughout this application various publications, patents, and/or patent applications are referenced. The disclosures of the publications, patents and/or patent applications are hereby incorporated by reference in their entireties into this application in order to more fully describe the state of the art to which this disclosure pertains.

Compositions

[0076] In some aspects, the pharmaceutical compositions of the present disclosure comprise a polynucleotide encoding a polypeptide comprising three units, *i.e.* a targeting unit, a dimerization unit, and an antigenic unit. Due to the dimerization unit, the polypeptide forms a dimeric protein called a Vaccibody™ as discussed herein.

[0077] In some embodiments, polynucleotides encoding the three units (*i.e.*, targeting unit, dimerization unit, and antigenic unit) are genetically engineered to be expressed from a single construct. When expressed *in vivo*, the polypeptides, *e.g.*, dimeric proteins, target antigen presenting cells (APCs); this targeting results in enhanced therapeutic potency compared to identical non-targeted antigens.

[0078] In some aspects, the present disclosure relates to pharmaceutical compositions, *e.g.*, vaccines, wherein the antigenic unit comprises antigenic subunits, wherein each subunit comprises a cancer neoepitope sequence or at least a part of a cancer neoepitope sequence, *e.g*., an epitope derived from calreticulin. The neoepitope sequence is obtained by sequencing tumor DNA or RNA and identifying tumor specific mutations representing neoantigens, *e.g*., at least one epitope derived from calreticulin frameshift mutation. Thereby, a personalized neoantigen vaccine may be obtained that specifically targets the identified tumor antigens.

Antigenic unit

[0079] In some aspects, an antigenic unit according to the disclosure comprises a plurality of tumor neoepitopes. In some embodiments, each neoepitope corresponds to at least one mutation identified in a tumor neoantigen. An example of tumor neoepitopes of the disclosure includes epitopes derived from the frameshift region of a calreticulin protein encoded by a sequence comprising a frameshift mutation.

[0080] In the antigenic unit, all but the last (*e.g.*, the most N terminal) of the tumor neoepitopes are arranged in antigenic subunits, wherein each subunit consists of a tumor neoepitope sequence and a linker, whereas the last subunit comprises a neoepitope only, *i.e.*, no such linker. Without wishing to be bound by theory, it is hypothesized that the separation of the tumor neoepitope sequences by the linker(s) causes each neoepitope to be presented in an optimal way to the immune system, ensuring therapeutic efficacy.

[0081] In some embodiments, the cancer neoepitope sequence, *e.g.*, an epitope derived from calreticulin comprising a frameshift mutation, comprises a length suitable for presentation by MHC I or MHC II. In some embodiments, the cancer neoepitope sequence, *e.g.*, an epitope derived from calreticulin frameshift mutation, is from 7 to 44 amino acids long. In some embodiments, the cancer neoepitope sequence, *e.g.*, an epitope derived from calreticulin comprising a frameshift mutation, is from 7 to 30 amino acids long. In some embodiments, the cancer neoepitope sequence, *e.g.*, an epitope derived from calreticulin comprising a frameshift mutation, has a length of from 7 to 10 amino acids, or from 13 to 30 amino acids. In some embodiments, the cancer neoepitope sequence, *e.g.*, an epitope derived from calreticulin comprising a frameshift mutation, is 9 amino acids. In some embodiments, the cancer neoepitope sequence, *e.g.*, an epitope derived from calreticulin comprising a frameshift mutation, is 15 amino acids.

[0082] Tumors can evade the immune system by shutting down expression of a mutated gene, for example, by an immune response directed towards the expression product of the mutated gene. In some embodiments, in order to reduce or prevent tumor avoidance of an immune response mediated by the polypeptides described herein, a plurality of different neoepitopes is used in the antigenic unit. Without wishing to be bound by theory, it is hypothesized that the more genes the tumor has to shut down, the less likely is it that the tumor is capable of immune evasion. Furthermore, the tumor may be heterogeneous in that not each and every neoantigen is expressed by all the tumor cells. Accordingly, a

combination of distinct neoepitopes can be used in the polypeptides described herein. Also, to ensure that all neoepitopes are loaded efficiently onto the same antigen presenting cell, the neoepitopes can be arranged as one amino acid chain instead of as discrete peptides. However, as described above, the object of the polypeptides is to activate T cells against the neoepitopes. T cells may be diluted in cases where too many neoepitopes or polynucleotides encoding the same are included in the same pharmaceutical composition, and therefore a balance can be struck to provide a pharmaceutical composition with an optimal number of neoepitopes in the antigenic unit.

[0083] In some aspects, the disclosure provides polypeptides comprising an antigenic unit comprising cancer neoepitopes, *e.g.*, one or more cancer neoepitopes. In some embodiments, at least 3 cancer neoepitopes, *e.g.*, epitopes derived from calreticulin comprising a frameshift mutation, are incorporated into the polypeptide, *e.g.*, at least 5 neoepitopes, at least 7 neoepitopes, at least 10 neoepitopes, or more. In some embodiments, the polypeptide comprises at least 10 neoepitopes. In some embodiments, the polypeptide comprises at least 15 neoepitopes. In some embodiments, the pharmaceutical composition according to the present disclosure comprises at least 20 neoepitopes.

[0084] In some embodiments, 15 to 50 cancer neoepitopes, *e.g.*, epitopes derived from calreticulin comprising a frameshift mutation, are included in the polypeptide, or polynucleotide encoding same. The incorporation of 15-50 cancer neoepitopes can elicit an efficient immune response without diluting the T cells when administered to a subject. In some embodiments, the polypeptide comprises an antigenic unit comprising between 15 to 40 neoepitopes, between 15 to 30 neoepitopes, between 15 to 25 neoepitopes, between 20 to 40 neoepitopes, between 20 to 30 neoepitopes, between 20 to 25 neoepitopes, or between 15 to 20 neoepitopes.

[0085] In some embodiments, the antigenic unit comprises one copy of each of a plurality of different cancer neoepitopes, *e.g.*, different epitopes derived from calreticulin comprising a frameshift mutation, so that when, *e.g.*, 10 neoepitopes are included in the polypeptide, a cell-mediated immune response against 10 different neoepitopes can be evoked when the pharmaceutical composition is administered to a subject.

[0086] The antigenic unit may comprise at least two copies of at least one neoepitope to strengthen the immune response to these neoepitopes. For example, the antigenic unit comprises multiple copies of each of two or more different neoepitopes (*e.g.*, two or more copies of a first neoepitope, two or more copies of a second neoepitope, two copies of a third

neoepitope, and the first, second and third neoepitopes are not the same). It is contemplated herein that for manufacturing and regulatory reasons, it may be an advantage to keep the length of plasmid and, *i.e.,* the antigenic unit, constant, and therefore it may be advantageous to include more than one copy of the same cancer neoepitope, *e.g.*, an epitope derived from calreticulin, *e.g.*, with a frameshift mutation, in the antigenic unit. As discussed above, it may be an advantage to keep the length of the antigenic unit constant, and therefore in some embodiments all the cancer neoepitope sequences have identical length.

[0087] The length of the antigenic unit can be determined by the length of the cancer neoepitopes, *e.g.*, epitope(s) derived from calreticulin comprising a frameshift mutation, and the number of cancer neoepitopes, *e.g.*, epitope(s) derived from calreticulin comprising a frameshift mutation, arranged in the antigenic unit, and the length accordingly may be from about 21 amino acids to 1500 amino acids, from about 30 amino acids to about 1000 amino acids, from about 50 amino acids to about 500 amino acids, from about 100 amino acids to about 400 amino acids, or from about 100 amino acids to about 300 amino acids.

[0088] In some embodiments, a neoepitope of the antigenic unit has a length of between 7 to 11 amino acids for MHC class I presentation. In some embodiments, the neoepitope has a length of 9 amino acids for MHC class I presentation.

[0089] In some embodiments, a neoepitope of the antigenic unit has a length of between 12 and 20 amino acids for MHC class II presentation. In some embodiments, the neoepitope has a length of 15 amino acids for MHC class II presentation.

[0090] When a cancer neoepitope of the antigenic unit, *e.g.*, epitope derived from calreticulin comprising a frameshift mutation, is short, such as a few amino acids long, the cancer neoepitope sequence comprises the neoepitope flanked at one or both sides by an amino acid sequence (for example, wild type calreticulin sequence that does not include the frameshifted sequence). In some embodiments, the cancer neoepitope is positioned essentially in the middle of a sequence, to ensure that the neoepitope is presented by the antigen presenting cells after processing. The amino acid sequences flanking one or both sides of the neoepitope may be the amino acid sequences flanking the neoepitope in the neoantigen, whereby the cancer neoepitope sequence is a true subsequence of the cancer neoantigen amino acid sequence. In some embodiments, the antigenic subunits are arranged in the order of more antigenic to less antigenic in the direction from the first linker towards the final neoepitope.

[0091] In some embodiments, a neoepitope of the antigenic unit, *e.g.*, an epitope derived from calreticulin, is derived from calreticulin comprising a frameshift mutation in exon 9 (*e.g.*, within the region of SEQ ID NO:1 from A352 to L417, *e.g.*, SEQ ID NO:171).

[0092] In some embodiments, the calreticulin frameshift mutation produces an amino acid sequence comprising a sequence from **Table 1** (excluding SEQ ID NO:1 or SEQ ID NO:171). In some embodiments, the calreticulin frameshift mutation comprises a sequence of any one of SEQ ID NOs: 2-3 or 141-158.

[0093] Example calreticulin sequences, including calreticulin frameshift sequences, are provided in **Table 1**.

Table 1. Calreticulin amino acid sequences

Peptide	**Amino Acid Sequence**
CALR sequence (wildtype)	MLLSVPLLLGLLGLAVAEPAVYFKEQFLDGDGWTSRWIESKHKSDFGKFVLSSGKFYGDEEKDK GLQTSQDARFYALSASFEPFSNKGQTLVVQFTVKHEQNIDCGGGYVKLFPNSLDQTDMHGDSEY NIMFGPDICGPGTKKVHVIFNYKGKNVLINKDIRCKDDEFTHLYTLIVRPDNTYEVKIDNSQVE SGSLEDDWDFLPPKKIKDPDASKPEDWDERAKIDDPTDSKPEDWDKPEHIPDPDAKKPEDWDEE MDGEWEPPVIQNPEYKGEWKPRQIDNPDYKGTWIHPEIDNPEYSPDPSIYAYDNFGVLGLDLWQ VKSGTIFDNFLITNDEAYAEEFGNETWGVTKAAEKQMKDKQDEEQRLKEEEEDKKRKEEEEAED KEDDEDKDEDEEDEEDKEEDEEEDVPGQAKDEL (SEQ ID NO:1)
CALR exon 9 (wildtype)	AAEKQMKDKQDEEQRLKEEEEDKKRKEEEEAEDKEDDEDKDEDEEDEEDKEEDEEEDVPGQAKD EL (SEQ ID NO:171)
CALR L367fs*46	TRRMMRTKMRMRRMRRTRRKMRRKMSPARPRTSCREACLQGWTEA (SEQ ID NO:2)
CALR E370fs*43	VMRTKMRMRRMRRTRRKMRRKMSPARPRTSCREACLQGWTEA (SEQ ID NO:141)
CALR E370fs*48	QRTRRMMRTKMRMRRMRRTRRKMRRKMSPARPRTSCREACLQGWTEA (SEQ ID NO:142)
CALR L367fs*48	QRTRRMMRTKMRMRRMRRTRRKMRRKMSPARPRTSCREACLQGWTEA (SEQ ID NO:143)
CALR L367fs*44	RMMRTKMRMRRMRRTRRKMRRKMSPARPRTSCREACLQGWTEA (SEQ ID NO:144)
CALR K368fs*51	RRRQRTRRMMRTKMRMRRMRRTRRKMRRKMSPARPRTSCREACLQGWTEA (SEQ ID NO:145)
CALR L367fs*52	RRRRQRTRRMMRTKMRMRRMRRTRRKMRRKMSPARPRTSCREACLQGWTEA (SEQ ID NO:146)
CALR R366fs*53	KRRRRQRTRRMMRTKMRMRRMRRTRRKMRRKMSPARPRTSCREACLQGWTEA (SEQ ID NO:147)
CALR E371fs*49	RQRTRRMMRTKMRMRRMRRTRRKMRRKMSPARPRTSCREACLQGWTEA (SEQ ID NO:148)
CALR K368fs*43	MMRTKMRMRRMRRTRRKMRRKMSPARPRTSCREACLQGWTEA (SEQ ID NO:149)
CALR E370fs*37	RMRRMRRTRRKMRRKMSPARPRTSCREACLQGWTEA (SEQ ID NO:150)
CALR D373fs*47	RTRRMMRTKMRMRRMRRTRRKMRRKMSPARPRTSCREACLQGWTEA (SEQ ID NO:151)
CALR K374fs*53	KRRRRQRTRRMMRTKMRMRRMRRTRRKMRRKMSPARPRTSCREACLQGWTEA (SEQ ID NO:152)
CALR E371fs*49	RTRRMMRTKMRMRRMRRTRRKMRRKMSPARPRTSCREACLQGWTEA (SEQ ID NO:153)

Peptide	Amino Acid Sequence
CALR K385fs*47	NCDRRMMRTKMRMRRMRRTRRKMRRKMSPARPRTSCREACLQGWTEA (SEQ ID NO:154)
CALR K385fs*47	LCRRMMRTKMRMRRMRRTRRKMRRKMSPARPRTSCREACLQGWTEA (SEQ ID NO:155)
CALR R376fs*55	LCKRRRRQRTRRMMRTKMRMRRMRRTRRKMRRKMSPARPRTSCREACLQGWTEA (SEQ ID NO:156)
CALR K385fs*47	SCRRMMRTKMRMRRMRRTRRKMRRKMSPARPRTSCREACLQGWTEA (SEQ ID NO:157)
CALR E381fs*48	DPCRRMMRTKMRMRRMRRTRRKMRRKMSPARPRTSCREACLQGWTEA (SEQ ID NO:158)
CALR Mutation Fragment 44-mer	RRMMRTKMRMRRMRRTRRKMRRKMSPARPRTSCREACLQGWTEA (SEQ ID NO:3)
CALR 9mer-1	RPRTSCREA (SEQ ID NO:4)
CALR 9mer-2	SPARPRTSC (SEQ ID NO:5)
CALR 9mer-3	RRMRRTRRK (SEQ ID NO:6)
CALR 9mer-4	RRMMRTKMR (SEQ ID NO:7)
CALR 9mer-7	RTKMRMRRM (SEQ ID NO:8)
CALR 9mer-8	MRMRRMRRT (SEQ ID NO:9)
CALR 9mer-9	RMMRTKMRM (SEQ ID NO:10)
CALR 9mer-10	RMRRTRRKM (SEQ ID NO:11)

[0094] In some embodiments, the calreticulin frameshift mutation produces a polypeptide comprising a sequence of SEQ ID NO:2. In some embodiments, the at least one epitope derived from calreticulin comprising a frameshift mutation has a length of about 7 to 44 amino acids from SEQ ID NO:2. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 7 to 11 amino acids from SEQ ID NO:2. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 9 amino acids from SEQ ID NO:2. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 15 amino acids from SEQ ID NO:2.

[0095] In some embodiments, the calreticulin frameshift mutation produces a polypeptide comprising a sequence of SEQ ID NO:3. In some embodiments, the at least one epitope derived from calreticulin comprising a frameshift mutation has a length of about 7 to 44 amino acids from SEQ ID NO:3. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 7 to 11 amino acids from SEQ ID NO:3. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 9

amino acids from SEQ ID NO:3. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 15 amino acids from SEQ ID NO:3.

[0096] An example, non-limiting nucleotide sequence encoding such polypeptide, *e.g.*, SEQ ID NO:3, comprises, consists of, or consists essentially of 5′-CGGAGAATGATGCGGACCAAGATGAGAATGAGGCGCATGCGGAGAACAAGGCGCAAGATGCGGAGAAAGATGAGCCCAGCCAGGCCTCGCACCTCCTGCAGGGAGGCATGTCTGCAGGGATGGACAGAGGCC-3′ (SEQ ID NO:183). In some embodiments, the nucleotide sequence comprises a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical to SEQ ID NO:183. In some embodiments, the nucleotide sequence comprises a sequence of SEQ ID NO:183 or a sequence with 1, 2, 3, 4 or 5 insertions, substitutions or deletions relative thereto.

[0097] In some embodiments, the calreticulin frameshift mutation produces a polypeptide comprising a sequence of SEQ ID NO:141. In some embodiments, the at least one epitope derived from calreticulin comprising a frameshift mutation has a length of about 7 to 44 amino acids and comprises SEQ ID NO:141 or a portion thereof. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 7 to 11 amino acids from SEQ ID NO:141. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 9 amino acids from SEQ ID NO:141. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 15 amino acids from SEQ ID NO:141.

[0098] In some embodiments, the calreticulin frameshift mutation produces a polypeptide comprising a sequence of SEQ ID NO:142. In some embodiments, the at least one epitope derived from calreticulin comprising a frameshift mutation has a length of about 7 to 44 amino acids from SEQ ID NO:142. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 7 to 11 amino acids from SEQ ID NO:142. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 9 amino acids from SEQ ID NO:142. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 15 amino acids from SEQ ID NO:142.

[0099] In some embodiments, the calreticulin frameshift mutation produces a polypeptide comprising a sequence of SEQ ID NO:143. In some embodiments, the at least one epitope derived from calreticulin comprising a frameshift mutation has a length of about 7 to 44 amino acids from SEQ ID NO:143. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 7 to 11 amino acids from SEQ ID NO:143. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 9

amino acids from SEQ ID NO:143. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 15 amino acids from SEQ ID NO:143.

[00100] In some embodiments, the calreticulin frameshift mutation produces a polypeptide comprising a sequence of SEQ ID NO:144. In some embodiments, the at least one epitope derived from calreticulin comprising a frameshift mutation has a length of about 7 to 44 amino acids and comprises SEQ ID NO:144 or a portion thereof. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 7 to 11 amino acids from SEQ ID NO:144. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 9 amino acids from SEQ ID NO:144. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 15 amino acids from SEQ ID NO:144.

[00101] In some embodiments, the calreticulin frameshift mutation produces a polypeptide comprising a sequence of SEQ ID NO:145. In some embodiments, the at least one epitope derived from calreticulin comprising a frameshift mutation has a length of about 7 to 44 amino acids from SEQ ID NO:145. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 7 to 11 amino acids from SEQ ID NO:145. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 9 amino acids from SEQ ID NO:145. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 15 amino acids from SEQ ID NO:145.

[00102] In some embodiments, the calreticulin frameshift mutation produces a polypeptide comprising a sequence of SEQ ID NO:146. In some embodiments, the at least one epitope derived from calreticulin comprising a frameshift mutation has a length of about 7 to 44 amino acids from SEQ ID NO:146. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 7 to 11 amino acids from SEQ ID NO:146. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 9 amino acids from SEQ ID NO:146. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 15 amino acids from SEQ ID NO:146.

[00103] In some embodiments, the calreticulin frameshift mutation produces a polypeptide comprising a sequence of SEQ ID NO:147. In some embodiments, the at least one epitope derived from calreticulin comprising a frameshift mutation has a length of about 7 to 44 amino acids from SEQ ID NO:147. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 7 to 11 amino acids from SEQ ID NO:147. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 9

amino acids from SEQ ID NO:147. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 15 amino acids from SEQ ID NO:147.

[00104] In some embodiments, the calreticulin frameshift mutation produces a polypeptide comprising a sequence of SEQ ID NO:148. In some embodiments, the at least one epitope derived from calreticulin comprising a frameshift mutation has a length of about 7 to 44 amino acids from SEQ ID NO:148. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 7 to 11 amino acids from SEQ ID NO:148. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 9 amino acids from SEQ ID NO:148. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 15 amino acids from SEQ ID NO:148.

[00105] In some embodiments, the calreticulin frameshift mutation produces a polypeptide comprising a sequence of SEQ ID NO:149. In some embodiments, the at least one epitope derived from calreticulin comprising a frameshift mutation has a length of about 7 to 44 amino acids and comprises SEQ ID NO:149 or a portion thereof. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 7 to 11 amino acids from SEQ ID NO:149. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 9 amino acids from SEQ ID NO:149. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 15 amino acids from SEQ ID NO:149.

[00106] In some embodiments, the calreticulin frameshift mutation produces a polypeptide comprising a sequence of SEQ ID NO:150. In some embodiments, the at least one epitope derived from calreticulin comprising a frameshift mutation has a length of about 7 to 44 amino acids and comprises SEQ ID NO:150 or a portion thereof. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 7 to 11 amino acids from SEQ ID NO:150. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 9 amino acids from SEQ ID NO:150. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 15 amino acids from SEQ ID NO:150.

[00107] In some embodiments, the calreticulin frameshift mutation produces a polypeptide comprising a sequence of SEQ ID NO:151. In some embodiments, the at least one epitope derived from calreticulin comprising a frameshift mutation has a length of about 7 to 44 amino acids from SEQ ID NO:151. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 7 to 11 amino acids from SEQ ID NO:151. In some

embodiments, the at least one epitope comprises, consists essentially of, or consists of 9 amino acids from SEQ ID NO:151. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 15 amino acids from SEQ ID NO:151.

[00108] In some embodiments, the calreticulin frameshift mutation produces a polypeptide comprising a sequence of SEQ ID NO:152. In some embodiments, the at least one epitope derived from calreticulin comprising a frameshift mutation has a length of about 7 to 44 amino acids from SEQ ID NO:152. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 7 to 11 amino acids from SEQ ID NO:152. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 9 amino acids from SEQ ID NO:152. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 15 amino acids from SEQ ID NO:152.

[00109] In some embodiments, the calreticulin frameshift mutation produces a polypeptide comprising a sequence of SEQ ID NO:153. In some embodiments, the at least one epitope derived from calreticulin comprising a frameshift mutation has a length of about 7 to 44 amino acids from SEQ ID NO:153. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 7 to 11 amino acids from SEQ ID NO:153. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 9 amino acids from SEQ ID NO:153. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 15 amino acids from SEQ ID NO:153.

[00110] In some embodiments, the calreticulin frameshift mutation produces a polypeptide comprising a sequence of SEQ ID NO:154. In some embodiments, the at least one epitope derived from calreticulin comprising a frameshift mutation has a length of about 7 to 44 amino acids from SEQ ID NO:154. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 7 to 11 amino acids from SEQ ID NO:154. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 9 amino acids from SEQ ID NO:154. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 15 amino acids from SEQ ID NO:154.

[00111] In some embodiments, the calreticulin frameshift mutation produces a polypeptide comprising a sequence of SEQ ID NO:155. In some embodiments, the at least one epitope derived from calreticulin comprising a frameshift mutation has a length of about 7 to 44 amino acids from SEQ ID NO:155. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 7 to 11 amino acids from SEQ ID NO:155. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 9

amino acids from SEQ ID NO:155. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 15 amino acids from SEQ ID NO:155.

[00112] In some embodiments, the calreticulin frameshift mutation produces a polypeptide comprising a sequence of SEQ ID NO:156. In some embodiments, the at least one epitope derived from calreticulin comprising a frameshift mutation has a length of about 7 to 44 amino acids from SEQ ID NO:156. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 7 to 11 amino acids from SEQ ID NO:156. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 9 amino acids from SEQ ID NO:156. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 15 amino acids from SEQ ID NO:156.

[00113] In some embodiments, the calreticulin frameshift mutation produces a polypeptide comprising a sequence of SEQ ID NO:157. In some embodiments, the at least one epitope derived from calreticulin comprising a frameshift mutation has a length of about 7 to 44 amino acids from SEQ ID NO:157. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 7 to 11 amino acids from SEQ ID NO:157. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 9 amino acids from SEQ ID NO:157. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 15 amino acids from SEQ ID NO:157.

[00114] In some embodiments, the calreticulin frameshift mutation produces a polypeptide comprising a sequence of SEQ ID NO:158. In some embodiments, the at least one epitope derived from calreticulin comprising a frameshift mutation has a length of about 7 to 44 amino acids from SEQ ID NO:158. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 7 to 11 amino acids from SEQ ID NO:158. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 9 amino acids from SEQ ID NO:158. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of 15 amino acids from SEQ ID NO:158.

[00115] In some embodiments, the at least one epitope derived from calreticulin has a length of 9 amino acids. In some embodiments, the at least one epitope comprises, consists essentially of, or consists of an amino acid sequence selected from the group consisting of SEQ ID NOs: 4-11.

[00116] In some embodiments, the antigenic unit comprises a plurality of epitopes, each epitope independently selected from the group consisting of SEQ ID NOs: 4-11. In some embodiments, the antigenic unit comprises two or more epitopes that do not comprise the

same sequence. In some embodiments, the antigenic unit comprises two or more epitopes that comprise the same sequence.

Linkers

[00117] A protein or polypeptide as disclosed herein may comprise one or more linkers. In some embodiments, a protein or polypeptide of the disclosure, or a polynucleotide encoding the same, comprises, from 5′ to 3′, a targeting unit, a dimerization unit, and an antigenic unit. In some embodiments, a linker connects one or more units. In some embodiments, a linker connects one or more subunits within a unit, such as the neoepitopes of the antigenic subunits described herein.

[00118] A linker may connect one or more cancer neoepitopes in an antigenic unit, *e.g.*, one or more epitope(s) derived from calreticulin comprising a frameshift mutation. A linker may connect an antigenic unit and a dimerization unit. A linker may connect a dimerization unit and a targeting unit.

[00119] A linker may be a chemical bond, *e.g.*, one or more covalent bonds or non-covalent bonds. In some embodiments, the linker is covalent. In some embodiments, the linker is non-covalent. In some embodiments, the linker comprises a peptide linker. In some embodiments, the peptide linker is between 2-30, 5-30, 10-30, 15-30, 20-30, 25-30, 2-25, 5-25, 10-25, 15-25, 20-25, 2-20, 5-20, 10-20, 15-20, 2-15, 5-15, 10-15, 2-10, 5-10, or 2-5 amino acids in length, or greater than or equal to 2, 5, 10, 15, 20, 25, or 30 amino acids in length (and optionally up to 50, 40, 30, 25, 20, 15, 10, or 5 amino acids in length).

[00120] A linker may comprise flexible, rigid, and/or cleavable linkers described herein. In some embodiments, the linker includes at least one glycine, alanine, and serine amino acid to provide for flexibility. In some embodiments, the linker is a hydrophobic linker, such as including a negatively charged sulfonate group, polyethylene glycol (PEG) group, or pyrophosphate diester group. In some embodiments, a linker is cleavable to selectively release a moiety (*e.g.*, a polypeptide) from a modulating agent, but is sufficiently stable to prevent premature cleavage.

[00121] As will be appreciated by one of skill in the art, commonly used flexible linkers have sequences consisting primarily of stretches of Gly and Ser residues (“GS” linker). Flexible linkers may be useful for joining domains/moieties that require a certain degree of movement or interaction and may include small, non-polar (*e.g.*, Gly) or polar (*e.g.*, Ser or Thr) amino acids. Incorporation of Ser or Thr can also maintain the stability of a linker in

aqueous solutions by forming hydrogen bonds with water molecules, and therefore reduce unfavorable interactions between a linker and moieties/domains.

[00122] Rigid linkers may be useful to keep a fixed distance between domains/moieties and to maintain their independent functions. Rigid linkers may also be useful when a spatial separation of domains is important for preserving the stability or bioactivity of one or more components in a fusion protein. Rigid linkers may have an alpha helix-structure or Pro-rich sequence, $(XP)_n$, with X designating any amino acid, preferably Ala, Lys, or Glu.

[00123] Cleavable linkers may release free functional domains *in vivo*. In some embodiments, linkers may be cleaved under specific conditions, such as in the presence of reducing reagents or proteases. *In vivo* cleavable linkers may utilize the reversible nature of a disulfide bond. One example of a cleavable linker is a linker that includes a thrombin-sensitive sequence (*e.g.*, PRS) between the two Cys residues. *In vitro* thrombin treatment of a fusion protein comprising a CPRSC (SEQ ID NO:172) sequence results in the cleavage of the thrombin-sensitive sequence, while a reversible disulfide linkage remains intact. Such linkers are known and described, *e.g.*, in Chen *et al*. 2013. Fusion Protein Linkers: Property, Design and Functionality. Adv Drug Deliv Rev. 65(10): 1357–1369. *In vivo* cleavage of linkers in fusion proteins may also be carried out by proteases that are expressed *in vivo* under certain conditions, in specific cells or tissues, or constrained within certain cellular compartments. Specificity of many proteases offers slower cleavage of the linker in constrained compartments. In some embodiments, the cleavable linker may be a self-cleaving peptide linker, *e.g.*, a T2A, P2A, E2A, or F2A peptide linker. Example sequences of self-cleaving peptides include EGRGSLLTCGDVEENPGP (T2A) (SEQ ID NO:173), ATNFSLLKQAGDVEENPGP (P2A) (SEQ ID NO:174), QCTNYALLKLAGDVESNPGP (E2A) (SEQ ID NO:175), and VKQTLNFDLLKLAGDVESNPGP (F2A) (SEQ ID NO:176). In some embodiments, the linker may comprise a "ribosome skipping" sequence, *e.g.*, a tPT2A linker.

[00124] Examples of molecules suitable for use in linkers described herein include a negatively charged sulfonate group; lipids, such as a poly (--CH_2--) hydrocarbon chains, such as polyethylene glycol (PEG) group, unsaturated variants thereof, hydroxylated variants thereof, amidated or otherwise N-containing variants thereof; noncarbon linkers; carbohydrate linkers; phosphodiester linkers, or other molecule capable of covalently linking two or more components of a protein or polypeptide. Non-covalent linkers are also included, such as hydrophobic lipid globules to which the polypeptide is linked, for example through a hydrophobic region of a polypeptide or a hydrophobic extension of a polypeptide, such as a

series of residues rich in leucine, isoleucine, valine, or perhaps also alanine, phenylalanine, or even tyrosine, methionine, glycine, or other hydrophobic residues. Components of a protein or polypeptide may be linked using charge-based chemistry, such that a positively charged component of a protein or polypeptide is linked to a negative charge of another component.

[00125] In some embodiments, in particular if the hydrophilicity/hydrophobicity varies greatly among the neoepitopes of the polypeptide, the most hydrophobic antigenic subunit(s) may be substantially positioned in the middle of the antigenic unit and the most hydrophilic antigenic subunit(s) positioned at the beginning and/or end of the antigenic unit. Alternatively, the neoepitopes may be arranged as alternating between a hydrophilic and a hydrophobic neoepitope. Furthermore, GC rich neoepitopes should be spaced so that GC clusters are avoided, *e.g.*, GC rich neoepitopes are spaced by at least one subunit. Neoepitope hydrophilicity/hydrophobicity can be determined from the neoepitope sequence by persons of ordinary skill in the art. In general, Leu, Ile, Phe, Trp, Val and Met are considered very hydrophobic; Cys, Tyr and Ala are considered hydrophobic; Thr, Glu, Gly, Ser, Gln, and Asp are considered neutral; and Arg, Lys, Asn, His, and Pro are considered hydrophilic.

[00126] The antigenic unit and the dimerization unit may be connected through a first linker. The first linker may further comprise a restriction site to facilitate the construction of the polynucleotide. In some embodiments, the first linker is GLGGL (SEQ ID NO:98) or GLSGL (SEQ ID NO:139).

[00127] In some embodiments, a protein or polypeptide of the disclosure comprises one or more of a second linker between neoepitopes within the antigenic unit. The number of second linkers is thus dependent on the number of neoepitopes present in the antigenic unit.

[00128] The second linker can be designed to be non-immunogenic, and may be a flexible linker, whereby the tumor neoepitopes, in spite of the high numbers of antigenic subunits present in the antigenic unit, are presented in an optimal manner to the T cells.

[00129] The second linker may be identical in all antigenic subunits. If, however, one or more of the neoepitopes comprises an amino acid motif similar to the linker, it may be an advantage to substitute the neighboring second linkers with a second linker of a different sequence. Also, if a neoepitope-second linker junction is predicted to constitute an epitope in itself, then a second linker of a different sequence might be used. In some embodiments, each second linker, *e.g.*, each linker between neoepitopes in the antigenic unit, comprises the same second linker sequence. In some embodiments, the antigenic unit comprises a plurality of

second linkers linking neoepitope sequences, and individual second linkers in the plurality are not the same.

[00130] The length of a linker (*e.g.*, flexible linker) can be "short," *e.g.*, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 or 12 amino acid residues, or "long," *e.g.*, at least 13 amino acid residues. In certain embodiments, a linker is 10-50, 10-40, 10-30, 10-25, 10-20, 15-50, 15-40, 15-30, 15-25, 15-20, 20-50, 20-40, 20-30, or 20-25 amino acid residues in length.

[00131] In certain embodiments, a linker comprises, consists essentially of, or consists of a (GS)n, (GGS)n, (GGGS)n (SEQ ID NO:179), (GGSG)n (SEQ ID NO:180), (GGSGG)n (SEQ ID NO:181), and (GGGGS)n (SEQ ID NO:182) sequence, wherein n is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20. In certain embodiments, a linker comprises, consists essentially of, or consists of an amino acid sequence selected from SEQ ID NOs:79-139, 165, or 169, as listed in **Table 2**.

[00132] A first linker as described herein may comprise, consist essentially of, or consist of an amino acid sequence selected from SEQ ID NOs:79-139, 165, or 169.

[00133] A second linker as described herein may comprise, consist essentially of, or consist of an amino acid sequence selected from SEQ ID NOs:79-139, 165, or 169.

Table 2. Example Linker Sequences

SEQ ID NO	Amino Acid Sequence	SEQ ID NO	Amino Acid Sequence	SEQ ID NO	Amino Acid Sequence
79	LGGGS	101	GGLSL	123	GGGSLGGGSS
80	GLGGS	102	GGLLG	124	LGGGSLGGGS
81	GGLGS	103	GLGSL	125	GLGGSGLGGS
82	GGGLS	104	LGLSS	126	GGLGSGGLGS
83	GGGGL	105	GLGLS	127	GGGLSGGGLS
84	LGGSG	106	GGLLS	128	GGGGLGGGGL
85	GLGSG	107	GLGSL	129	LGGSGLGGSG
86	GGLSG	108	GLGSL	130	GLGSGGLGSG
87	GGGLG	109	LGGGSGGGGS	131	GGLSGGGLSG
88	GGGSL	110	GLGGSGGGGS	132	GGGLGGGGLG
89	LGGSS	111	GGLGSGGGGS	133	GGGSLGGGSL
90	GLGSS	112	GGGLSGGGGS	134	LGGSSLGGSS
91	GGLSS	113	GGGGLGGGGS	135	GLGSSGLGSS
92	GGGLS	114	LGGSGGGGSG	136	GGLSSGGLSS
93	GGGSL	115	GLGSGGGGSG	137	GGGLSGGGLS
94	LGLGS	116	GGLSGGGGSG	138	GGGSLGGGSL
95	GLGLS	117	GGGLGGGGSG	139	GLSGL

SEQ ID NO	Amino Acid Sequence	SEQ ID NO	Amino Acid Sequence	SEQ ID NO	Amino Acid Sequence
96	GLLGS	118	GGGSLGGGSG	165	SGGGGSGGGG
97	LGGLS	119	LGGSSGGGSS	169	GGGSSGGGSG
98	GLGGL	120	GLGSSGGGSS		
99	LGLSG	121	GGLSSGGGSS		
100	GLLSG	122	GGGLSGGGSS		

[00134] Alternative linkers may be selected from GSAT (SEQ ID NO:140) linkers and SEG linkers, or variants thereof. GSAT (SEQ ID NO:140) linkers include, for example and without limitation, GGSAGGSGSGSSGGSSGASGTGTAGGTGSGSGTGSG (SEQ ID NO:166). SEG linkers include, for example and without limitation, GGSGGGSEGGGSEGGGSEGGGSEGGGSEGGGSGGGS (SEQ ID NO:167). Other suitable linkers, such as the T cell epitope linker ELKTPLGDTTHT (SEQ ID NO:168), are known in the art, and described, for example, in WO 2022/233851.

Targeting unit

[00135] Due to the targeting unit, the polypeptide or protein of the disclosure leads to attraction of dendritic cells (DCs), neutrophils, and other immune cells. Thus, the polypeptide or protein comprising the targeting unit can not only target the antigens to specific cells, but can also facilitate a response-amplifying effect (adjuvant effect) by recruiting specific immune cells to the administration site of the pharmaceutical composition. Without wishing to be bound by theory, it is believed that this mechanism is of importance in a clinical setting where patients can receive the pharmaceutical composition without any additional adjuvants, since the pharmaceutical composition can provide an adjuvant effect.

[00136] The term "targeting unit" as used herein can refer to a unit that delivers the polypeptide or protein with an antigen of the disclosure, *e.g.*, a neoepitope derived from calreticulin, to an antigen presenting cell for MHC class II-restricted presentation to CD4+ T cells, or for providing cross presentation to CD8+ T cells by MHC class I restriction. The targeting unit may be connected through the dimerization unit to the antigenic unit, wherein the latter is in either the COOH-terminal or the NH2-terminal end of the polypeptide or protein. In some embodiments, the antigenic unit is in the COOH-terminal end of the polypeptide or protein. The targeting unit is designed to target the polypeptide or protein of

the disclosure to surface molecules expressed on the relevant antigen presenting cells (APC), such as molecules expressed exclusively on subsets of dendritic cells (DC).

[00137] Examples of such target surface molecules on APCs include, for example and without limitation, human leukocyte antigen (HLA), cluster of differentiation 14 (CD 14), cluster of differentiation 40 (CD40), CLEC9A C-type lectin domain containing 9A (CLEC9A), chemokine receptors and Toll-like receptors (TLRs). HLA is a major histocompatibility complex (MHC) in humans.

[00138] Thus, the targeting unit may comprise or consist of an antibody-binding region, such as the antibody variable domains (VL and VH), with specificity for HLA, CD14, CD40, CLEC9A, chemokine receptors or Toll-like receptors.

[00139] In other embodiments, the targeting unit comprises, consists essentially of, or consists of a synthetic or natural ligand. Examples include a soluble CD40 ligand, natural ligands like chemokines, for example in their chemokine human forms, *e.g.*, chemokine ligand 5, also called C-C motif ligand 5 (CCL5 or RANTES), macrophage inflammatory protein alpha and its isoforms, including mouse CCL3 (MIP-la), and human isoforms hCCL3, hCCL3Ll, hCCL3L2 and hCCL3L3, chemokine ligand 4 (CCL4) and its isoform CCL4L, chemokine ligand 19 (CCL19), chemokine ligand 20 (CCL20), chemokine ligand 21 (CCL21), chemokine motif ligand 1 or 2 (XCL1 or XCL2) and bacterial antigens, *e.g.*, flagellin. In some embodiments, the targeting unit has affinity for an MHC class II protein. Thus, in some embodiments, the targeting unit comprises, consists essentially of, or consists of an antibody-binding region, such as the antibody variable domains (VL and VH) with specificity for MHC class II proteins, for example selected from the group consisting of anti-HLA-DP, anti-HLA-DR and anti-pan HLA class II.

[00140] In some embodiments, the targeting unit comprises, consists essentially of, or consists of chemokine human macrophage inflammatory protein alpha (human MIP-1α or MIP-1alpha (hMIP-1α), also known in the art as G0S19-1, LD78P, LD78alpha, SCYA3, G0/G1 switch regulatory protein 19-1, C-C Motif Chemokine 3, PAT 464.1, SIS-Beta, Small-inducible cytokine A3, or CCL3L1), and its isoforms, including mouse (CCL3 or MIP-1α), and human isoforms hCCL3, hCCL3Ll, hCCL3L2 and hCCL3L3, chemokine ligand 4 (CCL4) and its isoform CCL4L, chemokine ligand 19 (CCL19), chemokine ligand 20 (CCL20), chemokine ligand 21 (CCL21), which binds to their cognate receptors, for example, CCR1, CCR3 and/or CCR5. In some embodiments, the targeting unit comprises a MIP-1α chemokine that binds to a CCR5 receptor. In some embodiments, the MIP-1α

chemokine comprises an amino acid sequence of APLAADTPTACCFSYTSRQIPQNFIADYFETSSQCSKPSVIFLTKRGRQVCADPSEEWVQKYVSDLELSA (SEQ ID NO:159). In some embodiments, the MIP-1α chemokine consists essentially of an amino acid sequence of SEQ ID NO:159. It is contemplated herein that the MIP-1α chemokine, *e.g.*, SEQ ID NO:159, may comprise 1, 2, or 3 conservative substitutions, insertions, or deletions that do not affect function. An example, non-limiting nucleotide sequence encoding such MIP-1α chemokine, *e.g.*, SEQ ID NO:159, comprises, consists of, or consists essentially of 5′-GCACCACTTGCTGCTGACACGCCGACCGCCTGCTGCTTCAGCTACACCTCCCGACAGATTCCACAGAATTTCATAGCTGACTACTTTGAGACGAGCAGCCAGTGCTCCAAGCCCAGTGTCATCTTCCTAACCAAGAGAGGCCGGCAGGTCTGTGCTGACCCCAGTGAGGAGTGGGTCCAGAAATACGTCAGTGACCTGGAGCTGAGTGCC -3′ (SEQ ID NO:184). In some embodiments, the nucleotide sequence comprises a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical to SEQ ID NO:184. In some embodiments, the nucleotide sequence comprises a sequence of SEQ ID NO:184 or a sequence with 1, 2, 3, 4 or 5 insertions, substitutions or deletions relative thereto.

[00141] In some embodiments, the targeting unit has affinity for an MHC class II protein. In some embodiments, the nucleotide sequence encoding the targeting unit encodes antibody variable domains (VL and VH) with specificity for MHC class II proteins, selected from anti-HLA-DP, anti-HLA-DR, and anti-HLA-II antibodies. In some embodiments, the targeting unit has affinity for a surface molecule selected from CD40, TLR-2, TLR-4, and TLR-5. In some embodiments, the nucleotide sequence encoding the targeting unit encodes the antibody variable domains (VL and VH) with specificity for anti-CD40, anti-TLR-2, anti-TLR-4, and anti-TLR-5. In some embodiments, the nucleotide sequence encoding the targeting unit encodes flagellin.

[00142] In some embodiments, the targeting unit has affinity for a chemokine receptor selected from CCR1, CCR3, and CCR5. In some embodiments, the nucleotide sequence encoding the targeting unit encodes the chemokine hMIP-1α (LD78beta), which binds to cognate receptors, CCR1, CCR3, and CCR5 expressed on the cell surface of APCs.

[00143] The binding of the polypeptide/dimeric protein of the disclosure to its cognate receptors leads to internalization in the APC and degradation of the proteins into small peptides that are loaded onto MHC molecules and presented to CD4+ and CD8+ T cells to induce tumor specific immune responses. Once stimulated with activated CD4+ T cells, CD8+ T cells will target and kill tumor cells expressing the same neoantigens.

Dimerization unit

[00144] The term "dimerization unit" as used herein, can refer to a sequence of amino acids between the antigenic unit and the targeting unit. Thus, the dimerization unit serves to connect the antigenic unit and the targeting unit, and facilitates dimerization of two monomeric polypeptides into a dimeric protein. Furthermore, the dimerization unit also provides the flexibility in the polypeptide/dimeric protein to allow optimal binding of the targeting unit to the surface molecules on the antigen presenting cells (APCs), even if they are located at variable distances. The dimerization unit may be any unit that fulfils these requirements.

[00145] In some embodiments, the dimerization unit comprises a dimerization unit linker. In some embodiments, the dimerization unit linker is a glycine-serine rich linker, *e.g.*, GGGSSGGGSG (SEQ ID NO:169). In some embodiments, the dimerization unit comprises a glycine-serine rich dimerization unit linker, *e.g.*, GGGSSGGGSG (SEQ ID NO:169).

[00146] Accordingly, in some embodiments, the dimerization unit may comprise a hinge region and optionally another domain that facilitates dimerization, and the hinge region and the other domain may be connected through a linker, *e.g.*, a dimerization unit linker. The term "hinge region" refers to a peptide sequence of the dimeric protein that facilitates the dimerization. The hinge region functions as a flexible spacer between the units allowing the two targeting units to bind simultaneously to two target molecules on APCs, even if they are expressed with variable distances. The hinge region may be Ig derived, such as derived from IgG3. The hinge region may contribute to the dimerization through the formation of covalent bond(s), *e.g.,* disulfide bridge(s). Thus, in some embodiments the hinge region has the ability to form one or more covalent bonds. The covalent bond can be, for example, a disulfide bridge.

[00147] In some embodiments, the other domain that facilitates dimerization is an immunoglobulin domain, such as a carboxyterminal C domain, or a sequence that is substantially identical to the C domain or a variant thereof. In some embodiments, the other domain that facilitates dimerization is a carboxyterminal C domain derived from an IgG.

[00148] The immunoglobulin domain contributes to dimerization through non-covalent interactions, *e.g.,* hydrophobic interactions. For example, the immunoglobulin domain has the ability to form dimers *via* noncovalent interactions. In some embodiments, noncovalent interactions are hydrophobic interactions.

[00149] In some embodiments, the dimerization unit does not comprise a CH2 domain. In some embodiments, the dimerization unit comprises, consists essentially of, or consists of hinge exons h1 and h4 connected through a linker to a CH3 domain of human IgG3. In some embodiments the linker that connect the hinge region and another domain that facilitate dimerization, such as an immunoglobulin domain, is a $G_3S_2G_3SG$ (SEQ ID NO:169) linker.

[00150] It is contemplated herein that the dimerization unit may have any orientation with respect to the antigenic unit and the targeting unit. In some embodiments, the antigenic unit is in the COOH-terminal end of the dimerization unit with the targeting unit in the N-terminal end of the dimerization unit. In some embodiments, the antigenic unit is in the N-terminal end of the dimerization unit with the targeting unit in the COOH-terminal end of the dimerization unit. In some embodiments, the antigenic unit is in the COOH-end of the dimerization unit.

[00151] In some embodiments, the hinge exon h1 from IgG3 comprises, consists of, or consists essentially of ELKTPLGDTTHT (SEQ ID NO:160). In some embodiments, the hinge exon h4 from IgG3 comprises, consists of, or consists essentially of EPKSCDTPPPCPRCP (SEQ ID NO:161). In some embodiments, the human CH3 domain of IgG3 comprises, consists of, or consists essentially of GQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESSGQPENNYNTTPPMLDSDGSFFLYSKLTVDKS RWQQGNIFSCSVMHEALHNRFTQKSLSLSPGK (SEQ ID NO:162). It is contemplated herein that the human CH3 domain, *e.g.*, SEQ ID NO:162, may comprise 1, 2, or 3 conservative substitutions that do not affect function.

[00152] In some embodiments, the dimerization unit comprises, consists of, or consists essentially of ELKTPLGDTTHTEPKSCDTPPPCPRCPGGGSSGGGSGGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIA VEWESSGQPENNYNTTPPMLDSDGSFFLYSKLTVDKSRWQQGNIFSCSVMHEALHNRFTQKSLSLSPGK (SEQ ID NO:178). An example, non-limiting nucleotide sequence encoding such dimerization unit, *e.g.*, SEQ ID NO:178, comprises, consists of, or consists essentially of 5′ - GAGCTCAAAACCCCACTTGGTGACACAACTCACACAGAGCCCAAATCTTGTGACACACCTCCCCCGTGCCCAAGG TGCCCAGGCGGTGGAAGCAGCGGAGGTGGAAGTGGAGGACAGCCCCGAGAACCACAGGTGTACACCCTGCCCCCA TCCCGGGAGGAGATGACCAAGAACCAGGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTACCCCAGCGACATCGCC GTGGAGTGGGAGAGCAGCGGGCAGCCGGAGAACAACTACAACACCACGCCTCCCATGCTGGACTCCGACGGCTCC TTCTTCCTCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAGCAGGGGAACATCTTCTCATGCTCCGTGATG CATGAGGCTCTGCACAACCGCTTCACGCAGAAGAGCCTCTCCCTGTCTCCGGGTAAA - 3′ (SEQ ID NO:185). In some embodiments, the nucleotide sequence comprises a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical to SEQ

ID NO:185. In some embodiments, the nucleotide sequence comprises a sequence of SEQ ID NO:185 or a sequence with 1, 2, 3, 4 or 5 insertions, substitutions or deletions relative thereto.

Signal peptide

[00153] In some embodiments, the polynucleotide further comprises a nucleotide sequence encoding a signal peptide. The signal peptide is included to facilitate secretion of the polypeptide encoded by the polynucleotide of the disclosure in the cells transfected with said polynucleotide.

[00154] In some embodiments, a signal peptide comprises, consists of, or consists essentially of a signal peptide of human MIP-1α. In some embodiments, the signal peptide from human MIP-1α comprises, consists of, or consists essentially of MQVSTAALAVLLCTMALCNQVLS (SEQ ID NO:163).

[00155] An example, non-limiting nucleotide sequence encoding such human MIP-1α, *e.g.*, SEQ ID NO:163, comprises, consists of, or consists essentially of 5′-ATGCAGGTCTCCACTGCTGCCCTTGCCGTCCTCCTCTGCACCATGGCTCTCTGCAACCAGGTCCTCTCT- 3′ (SEQ ID NO:186). In some embodiments, the nucleotide sequence comprises a sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical to SEQ ID NO:186. In some embodiments, the nucleotide sequence comprises a sequence of SEQ ID NO:186 or a sequence with 1, 2, 3, 4 or 5 insertions, substitutions or deletions relative thereto.

[00156] In some embodiments, a dimeric protein or polypeptide of the disclosure comprises a signal peptide of human MIP-1α, an amino acid sequence of human MIP-1α, a hinge exon h1 from IgG3, a hinge exon h4 from IgG3, a dimerization unit linker, and human CH3 domain of IgG3. In some embodiments, a dimeric protein or polypeptide of the disclosure comprises

MQVSTAALAVLLCTMALCNQVLSAPLAADTPTACCFSYTSRQIPQNFIADYFETSSQCSKPSVIFLTKRGRQVCA
DPSEEWVQKYVSDLELSAELKTPLGDTTHTEPKSCDTPPPCPRCPGGGSSGGGSGGQPREPQVYTLPPSREEMTK
NQVSLTCLVKGFYPSDIAVEWESSGQPENNYNTTPPMLDSDGSFFLYSKLTVDKSRWQQGNIFSCSVMHEALHNR
FTQKSLSLSPGK (SEQ ID NO:170).

Polypeptides and dimeric proteins

[00157] The disclosure further relates to polypeptides comprising the antigenic units and targeting units described above. In some embodiments, the polypeptides comprise an

antigenic unit, a dimerization unit, and a targeting unit. In some embodiments, the polypeptides comprise an antigenic unit, a dimerization unit, a targeting unit and a signal peptide. In some embodiments, for example those embodiments wherein the polypeptides comprise a dimerization unit, the polypeptides form a dimer. The polypeptides may be expressed *in vitro* for production of a pharmaceutical composition comprising the polypeptides or proteins according to the disclosure, or the polypeptide may be expressed *in vivo* as a result of administration of a pharmaceutical composition comprising a polynucleotide encoding the polypeptide as defined above.

[00158] Due to the presence of the dimerization unit, dimeric proteins are formed when polypeptides comprising the dimerization unit are expressed. The dimeric protein may be a homodimer, *i.e.*, wherein the two polypeptide chains are identical and consequently comprise identical neoepitopes, or the dimeric protein may be a heterodimer comprising two different monomeric polypeptides encoded in the antigenic units. The latter may be relevant if the amount of neoepitopes exceeds an upper size limit for the antigenic unit.

[00159] In some embodiments, the protein is a dimeric protein. In some embodiments, the dimeric protein is a homodimeric protein. In some embodiments, the dimeric protein is a heterodimeric protein.

[00160] In some embodiments, the polypeptide comprises, consists essentially of, or consists of

MQVSTAALAVLLCTMALCNQVLSAPLAADTPTACCFSYTSRQIPQNFIADYFETSSQCSKPSVIFLTKRGRQVCA
DPSEEWVQKYVSDLELSAELKTPLGDTTHTEPKSCDTPPPCPRCPGGGSSGGGSGGQPREPQVYTLPPSREEMTK
NQVSLTCLVKGFYPSDIAVEWESSGQPENNYNTTPPMLDSDGSFFLYSKLTVDKSRWQQGNIFSCSVMHEALHNR
FTQKSLSLSPGKGLGGLRRMMRTKMRMRRMRRTRRKMRRKMSPARPRTSCREACLQGWTEA (SEQ ID NO:177).

[00161] In some embodiments, the dimeric protein comprises, consists essentially of, or consists of two polypeptides, each comprising

MQVSTAALAVLLCTMALCNQVLSAPLAADTPTACCFSYTSRQIPQNFIADYFETSSQCSKPSVIFLTKRGRQVCA
DPSEEWVQKYVSDLELSAELKTPLGDTTHTEPKSCDTPPPCPRCPGGGSSGGGSGGQPREPQVYTLPPSREEMTK
NQVSLTCLVKGFYPSDIAVEWESSGQPENNYNTTPPMLDSDGSFFLYSKLTVDKSRWQQGNIFSCSVMHEALHNR
FTQKSLSLSPGKGLGGLRRMMRTKMRMRRMRRTRRKMRRKMSPARPRTSCREACLQGWTEA (SEQ ID NO:177).

[00162] In some embodiments, the dimeric protein comprises, consists essentially of, or consists of two polypeptides, each consisting of

MQVSTAALAVLLCTMALCNQVLSAPLAADTPTACCFSYTSRQIPQNFIADYFETSSQCSKPSVIFLTKRGRQVCA
DPSEEWVQKYVSDLELSAELKTPLGDTTHTEPKSCDTPPPCPRCPGGGSSGGGSGGQPREPQVYTLPPSREEMTK

NQVSLTCLVKGFYPSDIAVEWESSGQPENNYNTTPPMLDSDGSFFLYSKLTVDKSRWQQGNIFSCSVMHEALHNR FTQKSLSLSPGKGLGGLRRMMRTKMRMRRMRRTRRKMRRKMSPARPRTSCREACLQGWTEA (SEQ ID NO:177).

Vectors

[00163] Polynucleotides encoding the proteins or polypeptides described herein, or fragments thereof, may be incorporated into a vector. Vectors, including those derived from retroviruses such as lentivirus, are suitable tools to achieve long-term gene transfer since they allow long-term, stable integration of a transgene, and its propagation in daughter cells. Examples of suitable vectors include expression vectors, replication vectors, probe generation vectors, and sequencing vectors. In some embodiments, an expression vector may be provided to a cell in the form of a viral vector. Viruses that are useful as vectors include, but are not limited to, retroviruses, adenoviruses, adeno-associated viruses, herpes viruses, and lentiviruses. In general, a suitable vector contains an origin of replication functional in at least one organism, a promoter sequence, convenient restriction endonuclease sites, and one or more selectable markers.

[00164] In some aspects, the disclosure provides a vector comprising a nucleotide sequence as defined above. In some embodiments, the vector allows for easy exchange of the various units described above, in particularly the antigenic unit. In some embodiments, the expression vector may be pUMVC4a vector (comprising a CMV promoter, a CMV intron, and a beta globin polyA signal) or vectors based on a NTC9385R (Nanoplasmid™) vector backbone. Suitable vectors are described in WO2017/118695, the contents of which are incorporated by reference herein. The antigenic unit may be exchanged with an antigenic unit cassette restricted by the Sfil restriction enzyme cassette where the 5' site is incorporated in the GLGGL (SEQ ID NO:98) or GLSGL (SEQ ID NO:139) linker, *e.g.*, first linker, and the 3' site is included after the stop codon in the vector. An example, non-limiting nucleotide sequence encoding such first linker, *e.g.*, GLGGL (SEQ ID NO:98) , comprises, consists of, or consists essentially of 5′ - GGCCTCGGTGGCCTG - 3′ (SEQ ID NO:187).

[00165] In some embodiments, the vector comprises a promoter sequence operably linked to a sequence encoding the protein, polypeptide or fragment thereof described supra. In some embodiments, the promoter is constitutive. In some embodiments, the promoter is tissue or cell type specific. In some embodiments, the promoter comprises a Cytomegalovirus (CMV) promoter. In some embodiments, the vector comprises one or more sequences to enhance expression of the protein, polypeptide or fragment thereof, including, but not limited to, a

sequence encoding an intron, polyadenylation signal, or Woodchuck hepatitis virus (WHV) posttranscriptional regulatory element (WPRE). Selection of suitable vector elements will be known to one of ordinary skill in the art.

Polynucleotides

[00166] The disclosure provides polynucleotides encoding the polypeptides described above. The polynucleotide may comprise a DNA nucleotide sequence or an RNA nucleotide sequence, such as genomic DNA, cDNA, and RNA sequences, either double stranded or single stranded. In some embodiments, the polynucleotide is optimized to the species to express the polypeptide according to the disclosure, *i.e.*, human codon optimized.

[00167] Provided herein are polynucleotides encoding a protein or polypeptide of the present disclosure. In some embodiments, the polynucleotide comprises a sequence encoding an antigenic unit and a sequence encoding a targeting unit. In some embodiments, the polynucleotide comprises a sequence encoding an antigenic unit, a dimerization unit, and a targeting unit. In some embodiments, the polynucleotide comprises a sequence encoding an antigenic subunit comprising one or more neoepitopes derived from calreticulin comprising a frameshift mutation, optionally wherein the neoepitopes are linked *via* linkers, a dimerization unit comprising an immunoglobulin constant domain, and a targeting domain comprising a MIP-1α sequence. In some embodiments, the polynucleotide comprises a sequence encoding an antigenic subunit comprising one or more neoepitopes each comprising 7 to 44 amino acids of any one of SEQ ID NOs:2, 3, or 141-158; a sequence encoding a dimerization unit comprising a hinge exon h1 from IgG3, a hinge exon h4 from IgG3, a dimerization unit linker, and human CH3 domain of IgG3, *e.g.*, SEQ ID NO:178; and a sequence encoding a targeting domain comprising a MIP-1α sequence, *e.g.*, SEQ ID NO:159. In some embodiments, the polynucleotide comprises a sequence encoding an antigenic subunit comprising one or more neoepitopes each comprising 7 to 44 amino acids of any one of SEQ ID NOs:2, 3, or 141-158; a sequence encoding a dimerization unit comprising a hinge exon h1 from IgG3, a hinge exon h4 from IgG3, a dimerization unit linker, and human CH3 domain of IgG3 of SEQ ID NO:170; and a sequence encoding a targeting domain comprising a MIP-1α sequence of SEQ ID NO:159. In some embodiments, the polynucleotide comprises a sequence encoding an antigenic unit, a sequence encoding a dimerization unit and a sequence encoding a targeting unit, wherein the sequences of the antigenic unit and the dimerization unit and/or the dimerization unit and the targeting unit are separated *via* sequences encoding linkers as described herein.

[00168] In some embodiments, a protein or polypeptide of the present disclosure may be provided to a subject *via* a pharmaceutical composition comprising a polynucleotide encoding the same, wherein the polynucleotide is associated with sufficient other sequences to achieve expression in a system of interest (*e.g.*, in a particular cell type, tissue, organism, *etc.*).

[00169] In some embodiments, the protein comprises, consists essentially of, or consists of MQVSTAALAVLLCTMALCNQVLSAPLAADTPTACCFSYTSRQIPQNFIADYFETSSQCSKPSVIFLTKRGRQVCA DPSEEWVQKYVSDLELSAELKTPLGDTTHTEPKSCDTPPPCPRCPGGGSSGGGSGGQPREPQVYTLPPSREEMTK NQVSLTCLVKGFYPSDIAVEWESSGQPENNYNTTPPMLDSDGSFFLYSKLTVDKSRWQQGNIFSCSVMHEALHNR FTQKSLSLSPGKGLGGLRRMMRTKMRMRRMRRTRRKMRRKMSPARPRTSCREACLQGWTEA (SEQ ID NO:177).

[00170] Provided herein are compositions comprising polynucleotides that encode a protein or polypeptide of the disclosure, or fragment thereof. In some embodiments, the protein is a dimeric protein. In some embodiments, the polynucleotides may be or may include DNA, RNA, or any other nucleic acid moiety or entity as described herein. The polynucleotides may be prepared by any technology described herein or otherwise available in the art (*e.g.*, synthesis, cloning, amplification, *in vitro* or *in vivo* transcription, *etc.*). In some embodiments, the polynucleotides that encode the protein or polypeptide, or fragment thereof, may be operationally associated with one or more replication, integration, and/or expression signals appropriate and/or sufficient to achieve integration, replication, and/or expression of the provided nucleic acid in a system of interest (*e.g.*, in a particular cell, tissue, organism, *etc.*).

[00171] In some embodiments, a composition for delivering a protein or polypeptide described herein is or comprises a vector, *e.g.*, a viral vector, comprising one or more polynucleotides encoding the protein or polypeptide, or one or more components of the protein as described herein. For example, when the protein is a dimer, the polynucleotides can encode one or both of the polypeptides of the dimer.

[00172] The disclosure provides compositions of polynucleotides that encode a protein or polypeptide of the disclosure, or fragments thereof. The polynucleotides may be or may include DNA, RNA, or any other nucleic acid moiety or entity as described herein, and may be prepared by any technology described herein or otherwise available in the art (*e.g.*, synthesis, cloning, amplification, *in vitro* synthesis, replication in host cell such as *E. coli*, or *in vivo* transcription, *etc.*). The polynucleotide sequence may include, for example and without limitation, DNA, RNA, modified oligonucleotides (*e.g.*, chemical modifications,

such as modifications that alter the backbone linkages, sugar molecules, and/or nucleic acid bases), and artificial nucleic acids. In some embodiments, the polynucleotide sequence includes, for example and without limitation, genomic DNA, cDNA, peptide nucleic acids (PNA) or peptide oligonucleotide conjugates, locked nucleic acids (LNA), bridged nucleic acids (BNA), polyamides, triplex forming oligonucleotides, modified DNA, antisense DNA oligonucleotides, tRNA, mRNA, rRNA, modified RNA, miRNA, gRNA, and siRNA or other RNA or DNA molecules. In some embodiments, provided polynucleotides encoding a protein or polypeptide, or polypeptide fragments thereof, may be operationally associated with one or more replication, integration, and/or expression signals appropriate and/or sufficient to achieve integration, replication, and/or expression of the provided nucleic acid in a system of interest (*e.g.*, in a particular cell, tissue, organism, *etc.*).

Host cells

[00173] The present disclosure is further directed, in part, to cells comprising the proteins or polypeptides described herein, or fragments thereof, as well as polynucleotides or vectors encoding same. Any cell, *e.g.*, cell line, *e.g.*, a cell line suitable for expression of a recombinant polypeptide, polynucleotide or vector known to one of skill in the art is suitable. The selection of an appropriate cell will be within the knowledge of one of ordinary skill in the art.

[00174] In some embodiments, a cell, *e.g.*, cell line, may be used to express a protein or polypeptide, *e.g.*, Vaccibody™, described herein. In some embodiments, a cell, *e.g.*, cell line, may be used to express or amplify a polynucleotide, *e.g.*, a vector, encoding a protein or polypeptide, *e.g.*, Vaccibody™, described herein. In some embodiments, a cell comprises a polypeptide encoding a protein or polypeptide, *e.g.*, Vaccibody™, described herein.

[00175] The disclosure also relates to a host cell comprising a polynucleotide sequence as defined above, or comprising a vector as defined above for expression of the polypeptide according to the disclosure.

[00176] Suitable host cells include prokaryotes (for example, *E. coli*), yeast, insect or other eukaryotic host cells known in the art.

Pharmaceutical Compositions

[00177] The present disclosure is further directed, in part, to pharmaceutical compositions comprising one or more polynucleotides of the disclosure, *e.g.*, a Vaccibody™ comprising at least one epitope derived from calreticulin, as described herein.

[00178] The present disclosure is further directed, in part, to pharmaceutical compositions comprising one or more proteins or polypeptides of the disclosure as described herein.

[00179] In some embodiments, the pharmaceutical composition comprises one or more pharmaceutically acceptable adjuvants. Pharmaceutically acceptable adjuvants include, but are not limited to poly-ICLC, 1018 ISS, aluminum salts, Amplivax, AS 15, BCG, CP-870,893, CpG7909, CyaA, dSLIM, GM-CSF, IC30, IC31, Imiquimod, ImuFactEVl P321, IS Patch, ISS, ISCOMATRIX, Juvlmmune, LipoVac, MF59, monophosphoryl lipid A, Montanide IMS 1312, Montanide ISA 206, Montanide ISA 50V, Montanide ISA-51, OK-432, OM-174, OM-197-MP-EC, ONTAK, PLGA microparticles, resiquimod, SRL172, virosomes and other virus-like particles, YF-17D, VEGF trap, R848, beta- glucan, Pam3Cys, Aquila's QS21 stimulon, vadimezan, and/or AsA404 (DMXAA). However, due to the presence of the targeting units, the composition can be administered without additional adjuvant. Thus, in some embodiments, the pharmaceutical composition does not comprise an adjuvant.

[00180] In some embodiments, for example those embodiments wherein the polynucleotide, polypeptide or protein is encapsulated in a nanoparticle, *e.g.*, a lipid nanoparticle, the pharmaceutical composition comprises a pharmaceutically acceptable amphiphilic block co-polymer. In some embodiments, the amphiphilic block co-polymer comprises blocks of poly(ethylene oxide) and polypropylene oxide).

[00181] An "amphiphilic block co-polymer" as used herein is a linear or branched co-polymer comprising or consisting of blocks of poly(ethylene oxide) ("PEO") and blocks of polypropylene oxide) ("PPO"). Typical examples of useful PEO-PPO amphiphilic block co-polymers have the general structures PEO-PPO-PEO (poloxamers), PPO PEO PPO, (PEO PPO-) 4ED (a poloxamine), and (PPO PEO-)4ED (a reverse poloxamine), where "ED" is an ethylenediaminyl group.

[00182] A "poloxamer" is a linear amphiphilic block co-polymer constituted by one block of poly(ethylene oxide) coupled to one block of polypropylene oxide) coupled to one block of PEO, *i.e.*, a structure of the formula EOa-POb-EOa, where EO is ethylene oxide, PO is propylene oxide, a is an integer from 2 to 130, and b is an integer from 15 to 67. Poloxamers are conventionally named by using a 3-digit identifier, where the first 2 digits multiplied by 100 provides the approximate molecular mass of the PPO content, and where the last digit multiplied by 10 indicates the approximate percentage of PEO content. For example, "Poloxamer 188" refers to a polymer comprising a PPO block of a molecular weight of about

1800 (corresponding to b being about 31 PPO) and approximately 80% (w/w) of PEO (corresponding to a being about 82). However, the values are known to vary to some degree, and commercial products such as the research grade Lutrol® F68 and the clinical grade Kolliphor®.

[00183] Poloxamer 188 (P188) exhibits a large variation in molecular weight (between 7,680 and 9,510) and the values for a and b provided for these particular products are indicated to be approximately 79 and 28, respectively. This reflects the heterogeneous nature of the block co-polymers, meaning that the values of a and b are averages found in a final formulation.

[00184] A "poloxamine" or "sequential poloxamine" (commercially available under the trade name of Tetronic®) is an X-shaped block co-polymers that bears four PEO-PPO arms connected to a central ethylenediamine moiety via bonds between the free OH groups comprised in the PEO-PPO-arms and the primary amine groups in ethylenediamine moiety. Reverse poloxamines are likewise X- shaped block co-polymers that bear four PPO-PEO arms connected to a central ethylenediamine moiety via bonds between the free OH groups comprised in the PPO-PEO arms and the primary amine groups in ethylenediamine.

[00185] Example amphiphilic block co-polymers are poloxamers or poloxamines, for example and without limitation, poloxamer 407 and 188, in particular poloxamer 188. Poloxamines may be sequential poloxamines of formula (PEO-PPO)4-ED. Poloxamines may be marketed under the registered trademarks Tetronic® 904, 704, and 304, respectively. The characteristics of these poloxamines are as follows: Tetronic® 904 has a total average molecular weight of 6700, a total average weight of PPO units of 4020, and a PEO percentage of about 40%. Tetronic® 704 has a total average molecular weight of 5500, a total average weight of PPO units of 3300, and a PEO percentage of about 40%; and Tetronic® 304 has a total average molecular weight of 1650, a total average weight of PPO units of 990, and a PEO percentage of about 40%.

[00186] In some embodiments, the composition comprises the amphiphilic block copolymer in an amount of from 0.2% weight per volume (w/v) to 20% w/v, such as from 0.2% w/v to 18% w/v, 0.2% w/v to 16% w/v, 0.2% w/v to 14% w/v, 0.2% w/v to 12% w/v, 0.2% w/v to 10% w/v, 0.2% w/v to 8% w/v, 0.2% w/v to 6% w/v, 0.2% w/v to 4% w/v, 0.4% w/v to 18% w/v, 0.6% w/v to 18% w/v, 0.8% w/v to 18% w/v, 1% w/v to 18% w/v, 2% w/v to 18% w/v, 1% w/v to 5% w/v, or 2% w/v to 4% w/v. In some embodiments, amounts range of from 0.5% w/v to 5% w/v. In other embodiments, the composition comprises the

amphiphilic block co-polymer in an amount of from 2% w/v to 5% w/v, such as about 3% w/v. For pharmaceutical compositions comprising the polynucleotide or vector, the pharmaceutical compositions may comprise molecules that ease transfection of cells.

[00187] In some embodiments, the pharmaceutical composition comprises a polynucleotide, *e.g.*, comprised in a vector such as apolycistronic vector, and is administered by intramuscular or intradermal injection.

[00188] The pharmaceutical composition of the disclosure may comprise the polynucleotide in a range of from 0.1 to 10 mg per unit dosage, *e.g.*, about 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 or 1 mg or *e.g.*, 2, 3, 4, 5, 6, 7, 8, 9 or 10 mg.

[00189] The pharmaceutical composition of the disclosure typically comprises the polypeptide or protein in the range of from 5 pg to 5 mg.

[00190] The amount of polynucleotide, polypeptide, or protein may vary depending on whether the pharmaceutical composition is administered for prophylactic, to reduce the likelihood of disease, or for therapeutic treatment, the severity of the disease in individuals which are infected, and on parameters like the age, weight, gender, medical history and pre-existing conditions.

[00191] In some embodiments, pharmaceutical compositions may be specially formulated for administration in solid or liquid form, including those adapted for the following: oral administration, for example, drenches (aqueous or non-aqueous solutions or suspensions), tablets, *e.g.*, those targeted for buccal, sublingual, and systemic absorption, boluses, powders, granules, pastes for application to the tongue; parenteral administration, for example, by subcutaneous, intramuscular, intravenous or epidural injection as, for example, a sterile solution or suspension, or sustained-release formulation; topical application, for example, as a cream, ointment, or a controlled-release patch or spray applied to the skin, lungs, or oral cavity; intravaginally or intrarectally, for example, as a pessary, cream, or foam; sublingually; ocularly; trans-dermally; or nasally, pulmonary, and/or to other mucosal surfaces.

[00192] Pharmaceutically acceptable carriers are involved in carrying or transporting the subject compound from one organ, or portion of the body, to another organ, or portion of the body. Each carrier must be “acceptable” in the sense of being compatible with the other ingredients of the formulation and not injurious to the patient. In some embodiments, materials which can serve as pharmaceutically-acceptable carriers include: sugars, such as lactose, glucose and sucrose; starches, such as com starch and potato starch; cellulose, and its derivatives, such as sodium carboxymethyl cellulose, ethyl cellulose and cellulose acetate;

powdered tragacanth; malt; gelatin; talc; excipients, such as cocoa butter and suppository waxes; oils, such as peanut oil, cottonseed oil, safflower oil, sesame oil, olive oil, corn oil and soybean oil; glycols, such as propylene glycol; polyols, such as glycerin, sorbitol, mannitol and polyethylene glycol; esters, such as ethyl oleate and ethyl laurate; agar; buffering agents, such as magnesium hydroxide and aluminum hydroxide; alginic acid; pyrogen-free water; isotonic saline; Ringer's solution; ethyl alcohol; pH buffered solutions; polyesters, polycarbonates and/or polyanhydrides; and other non-toxic compatible substances employed in pharmaceutical formulations.

[00193] Pharmaceutically acceptable salts include, but are not limited to, nontoxic acid addition salts, which are salts of an amino group formed with inorganic acids such as hydrochloric acid, hydrobromic acid, phosphoric acid, sulfuric acid and perchloric acid or with organic acids such as acetic acid, maleic acid, tartaric acid, citric acid, succinic acid, or malonic acid or by using other methods used in the art such as ion exchange. In some embodiments, pharmaceutically acceptable salts include, but are not limited to, adipate, alginate, ascorbate, aspartate, benzenesulfonate, benzoate, bisulfate, borate, butyrate, camphorate, camphorsulfonate, citrate, cyclopentanepropionate, digluconate, dodecylsulfate, ethanesulfonate, formate, fumarate, glucoheptonate, glycerophosphate, gluconate, hemisulfate, heptanoate, hexanoate, hydroiodide, 2-hydroxy-ethanesulfonate, lactobionate, lactate, laurate, lauryl sulfate, malate, maleate, malonate, methanesulfonate, 2-naphthalenesulfonate, nicotinate, nitrate, oleate, oxalate, palmitate, pamoate, pectinate, persulfate, 3-phenylpropionate, phosphate, picrate, pivalate, propionate, stearate, succinate, sulfate, tartrate, thiocyanate, toluenesulfonate, undecanoate, valerate salts, and the like. Representative alkali or alkaline earth metal salts include sodium, lithium, potassium, calcium, magnesium, and the like. In some embodiments, pharmaceutically acceptable salts include, when appropriate, nontoxic ammonium, quaternary ammonium, and amine cations formed using counterions such as halide, hydroxide, carboxylate, sulfate, phosphate, nitrate, alkyl having from 1 to 6 carbon atoms, sulfonate, and aryl sulfonate.

[00194] In various embodiments, the present disclosure provides pharmaceutical compositions described herein with a pharmaceutically acceptable excipient. Pharmaceutically acceptable excipient includes an excipient that is useful in preparing a pharmaceutical composition that is generally safe, non-toxic, and desirable, and includes excipients that are acceptable for veterinary use as well as for human pharmaceutical use.

Such excipients may be solid, liquid, semisolid, or, in the case of an aerosol composition, gaseous.

[00195] Pharmaceutical preparations may be made following conventional techniques of pharmacy involving milling, mixing, granulation, and compressing, when necessary, for tablet forms; or milling, mixing, and filling for hard gelatin capsule forms. When a liquid carrier is used, a preparation can be in the form of a syrup, elixir, emulsion or an aqueous or non-aqueous solution or suspension. Such a liquid formulation may be administered directly per os.

[00196] In some embodiments, pharmaceutical compositions may be formulated for delivery to a cell and/or to a subject *via* any route of administration. Modes of administration to a subject may include injection, infusion, inhalation, intranasal, intraocular, topical delivery, inter-cannular delivery, or ingestion. Injection includes, without limitation, intravenous, intramuscular, intra-arterial, intrathecal, intraventricular, intracapsular, intra-orbital, intracardiac, intradermal, intraperitoneal, transtracheal, subcutaneous, subcuticular, intraarticular, sub capsular, subarachnoid, intraspinal, intra-cerebrospinal, and intra-stemal injection and infusion. In some embodiments, administration is intramuscular. Intramuscular administration can be performed with a needle, or it can be needle-free. In some embodiments, intramuscular injection is performed by a needle-free hypodermic injector system, *e.g.*, a PharmaJet™. In some embodiments, administration includes aerosol inhalation, *e.g.*, with nebulization. In some embodiments, administration is systemic (*e.g.*, oral, rectal, nasal, sublingual, buccal, or parenteral), enteral (*e.g.*, system-wide effect, but delivered through the gastrointestinal tract), or local (*e.g.*, local application on the skin, or intravitreal injection). In some embodiments, one or more compositions is administered systemically. In some embodiments, administration is non-parenteral and a therapeutic is a parenteral therapeutic. In some embodiments, administration may be bronchial (*e.g.*, by bronchial instillation), buccal, dermal (which may be or comprise, for example, one or more of topical to the dermis, intradermal, inter-dermal, transdermal, *etc.*), enteral, intra-arterial, intradermal, intragastric, intramedullary, intramuscular, intranasal, intraperitoneal, intrathecal, intravenous, intraventricular, within a specific organ (*e.g.* intrahepatic), mucosal, nasal, oral, rectal, subcutaneous, sublingual, topical, tracheal (*e.g.*, by intratracheal instillation), vaginal, vitreal, *etc.*

[00197] In some embodiments, administration may be a single dose. In some embodiments, administration may involve dosing that is intermittent (*e.g.*, a plurality of doses

separated in time) and/or periodic (*e.g.*, individual doses separated by a common period of time) dosing. In some embodiments, administration may involve continuous dosing (*e.g.*, perfusion) for at least a selected period of time. In some embodiments, six, eight, ten, 12, 15 or 20 or more administrations may be given to the subject during one treatment or over a period of time as a treatment regimen.

[00198] In some embodiments, administrations may be given as needed, *e.g.*, for as long as symptoms associated with the disease, disorder or condition persist. In some embodiments, repeated administrations may be indicated for the remainder of the subject's life. Treatment periods may vary and could be, *e.g.*, one day, two days, three days, one week, two weeks, one month, two months, three months, six months, a year, or longer.

[00199] In some embodiments, administration comprises a series of doses, *e.g.*, a loading dose followed by a maintenance dose. A loading dose can be administered to achieve therapeutically effective levels of a pharmaceutical composition of the disclosure, with administration being at higher dose or higher frequency than a maintenance dose. In some embodiments, the loading dose is administered more frequently than the maintenance dose. In some embodiments, the loading dose is administered at higher dose level than the maintenance dose. In some embodiments, the loading dose is administered at an equal dose level to the maintenance dose. . In some embodiments, a loading dose is administered once, twice, three times, four times, five times, or more than five times. In some embodiments, a maintenance dose is administered once, twice, three times, four times, five times, or more than five times.

[00200] In some embodiments, the loading dose and maintenance dose are about 3 mg, about 6 mg, or about 9 mg. In some embodiments, the loading dose and maintenance dose are both about 3 mg.

[00201] In some embodiments, the loading dose is administered weekly (QW). In some embodiments, the loading dose is administered biweekly (Q2W). In some embodiments, the loading dose is administered every 3 weeks (Q3W). In some embodiments, the loading dose is administered monthly or every 4 weeks (Q4W). In some embodiments, the maintenance dose is administered biweekly (Q2W). In some embodiments, the maintenance dose is administered every 3 weeks (Q3W). In some embodiments, the maintenance dose is administered monthly or every 4 weeks (Q4W). In some embodiments, the maintenance dose is administered every 6 weeks (Q6W). In some embodiments, the maintenance dose is administered every 12 weeks (Q12W).

[00202] In some embodiments, the loading dose is administered bi-weekly, *e.g.*, for five doses, and the maintenance dose is administered every 6 weeks. In some embodiments, administration is at about 3 mg, about 6 mg, or about 9 mg, by intramuscular injection, *e.g.*, needle-free injection.

[00203] In some embodiments, the loading dose is administered every 3 weeks, *e.g.*, for four doses, and the maintenance dose is administered every 6 weeks. In some embodiments, administration is at about 3 mg, about 6 mg, or about 9 mg, by intramuscular injection, *e.g.*, needle-free injection.

[00204] In some embodiments, the loading dose is administered every 3 weeks , *e.g.*, for four doses, and the maintenance dose is administered every 12 weeks. In some embodiments, administration is at about 3 mg, about 6 mg, or about 9 mg, by intramuscular injection, *e.g.*, needle-free injection.

Methods for preparing pharmaceutical compositions

[00205] Pharmaceutical composition of the present disclosure comprising one or more polynucleotides of the disclosure, *e.g.*, a calreticulin Vaccibody™ as described herein, can comprise polynucleotides that encode neoantigens as identified in the tumor of a subject. Accordingly, the pharmaceutical compositions described herein target the specific mutated calreticulin proteins in the patient's tumor.

[00206] The disclosure relates in part to methods for preparing a pharmaceutical composition comprising an immunologically effective amount of a polynucleotide of the disclosure, or a protein or polypeptide encoded thereby, *in vitro*. Such *in vitro* synthesis may be carried out by any suitable method known to the person skilled in the art, such as through peptide synthesis or expression of the polypeptide in any of a variety of expressions systems followed by purification.

[00207] Methods of producing pharmaceutical compositions comprising the polypeptides or proteins of the disclosure include, for example and without limitation, a) transfecting a cell population with a polynucleotide disclosed herein; b) culturing the cell population; c) collecting and purifying the protein or the polypeptide expressed by the cell population, and d) mixing the protein or polypeptide obtained under step c) with a pharmaceutically acceptable carrier, thereby obtaining the pharmaceutical composition.

[00208] In some embodiments, a buffer may be added to the pharmaceutical composition. In some embodiments, a buffer may be added to the pharmaceutical composition.

[00209] In some embodiments, an adjuvant may be added to the pharmaceutical composition.

[00210] Purification may be carried out according to any suitable method known in the art, such as chromatography, centrifugation, or differential solubility.

[00211] Methods of producing pharmaceutical compositions comprising the polynucleotides of the disclosure include, for example and without limitation a) preparing the polynucleotide of the disclosure by any suitable methods known in the art; and b) mixing the polynucleotide obtained under step a) with a pharmaceutically acceptable carrier, thereby obtaining the pharmaceutical composition.

[00212] The polynucleotide may be prepared by any suitable method known to the skilled person. For example and without limitation, the polynucleotide may be prepared by chemical synthesis using an oligonucleotide synthesizer. Alternatively, the polynucleotide may be produced by a host cell such as *E. coli* and purified using any suitable methods known in the art.

[00213] In some embodiments, for example those embodiments with smaller nucleotide sequences, *e.g.*, nucleotide sequences encoding the targeting unit, the dimerization unit, the subunits of the antigenic unit, such sequences may be synthesized individually and then ligated together, to produce the final polynucleotide. In some embodiments, the final polynucleotide is introduced into a vector backbone.

[00214] In some embodiments, wherein the pharmaceutical composition is a personalized pharmaceutical composition, the methods described *supra* are preceded by a method of identifying the neoepitopes to be included in the polynucleotide.

[00215] This method of identifying the neoepitopes can include the steps of a) sequencing the genome or exome of a tumor, or otherwise determining the status of a calreticulin mutation in a patient's tumor, b) identifying the calreticulin neoepitopes from the tumor, and c) selecting the neoepitope or a combination thereof based on predicted antigenicity.

[00216] In some embodiments, sequencing the tumor may be by through any suitable method, such as by obtaining a biopsy of the tumor, or by excision of the tumor, or from collection of any suitable body fluid, such as a blood sample or a urine sample.

[00217] In some embodiments, the pharmaceutical composition may further comprise a pharmaceutically acceptable carrier, diluent, adjuvant, or buffer.

[00218] In some embodiments, wherein the pharmaceutical composition comprises polynucleotides, carriers may include molecules that ease transfection of cells. Adjuvants

may include plasmids comprising nucleotide sequences encoding chemokines or cytokines in order to enhance the immune response.

[00219] A pharmaceutical composition of the disclosure may be formulated into any suitable formulation, such as a liquid formulation for intradermal or intramuscular injection.

Administration

[00220] A pharmaceutical composition of the disclosure may be administered in any suitable route of administration for either a polypeptide/protein vaccine or a polynucleotide vaccine, such as administration by injection intradermally, intramuscular, subcutaneously, or by mucosal or epithelial application, such as intranasally, orally, enteral or to the bladder.

[00221] In some embodiments, wherein the pharmaceutical composition comprises polynucleotides, administration may be intramuscular or intradermal. In some embodiments, administration is intramuscular. Intramuscular administration can be performed with a needle, or it can be needle-free. In some embodiments, intramuscular injection is performed by a needle-free hypodermic injector system, *e.g.*, a PharmaJet™.

[00222] In some embodiments, the pharmaceutical composition is administered by intranodal injection. As used herein, the term "intranodal injection" means that the pharmaceutical composition is injected into the lymph nodes.

[00223] Pharmaceutical compositions can be administered by injection, infusion, inhalation, intranasal, intraocular, topical delivery, inter-cannular delivery, or ingestion. In some embodiments, administration includes aerosol inhalation, *e.g.*, with nebulization. In some embodiments, administration is systemic, enteral, or local. In some embodiments, administration is systemic. In some embodiments, administration is non-parenteral. In some embodiments, administration is bronchial, buccal, dermal, enteral, intra-arterial, intradermal, intragastric, intramedullary, intramuscular, intranasal, intraperitoneal, intrathecal, intravenous, intraventricular, within a specific organ (*e.g.,* intrahepatic), mucosal, nasal, oral, rectal, subcutaneous, sublingual, topical, tracheal, vaginal, vitreal, *etc.*

[00224] In some embodiments, the pharmaceutical composition is administered as a single dose. In some embodiments, the pharmaceutical composition is administered using intermittent and/or periodic dosing. In some embodiments, administration may involve continuous dosing (*e.g.*, perfusion) for at least a selected period of time. In some embodiments, six, eight, ten, 12, 15, or 20 or more administrations may be given to the subject during one treatment or over a period of time as a treatment regimen.

[00225] Pharmaceutical compositions according to the present disclosure may be delivered in a therapeutically effective amount. A precise therapeutically effective amount is an amount of a composition that will yield the most effective results in terms of efficacy of treatment in a given subject. This amount will vary depending upon a variety of factors, including but not limited to characteristics of the therapeutic compound (including activity, pharmacokinetics, pharmacodynamics, and bioavailability), physiological condition of a subject (including age, sex, disease type and stage, general physical condition, responsiveness to a given dosage, and type of medication), nature of a pharmaceutically acceptable carrier or carriers in a formulation, and/or route of administration.

Methods of Use

Treatment of Cancer

[00226] In some aspects, the polynucleotides, polypeptides, and proteins of the disclosure are for methods of use in the treatment of cancer. In some embodiments, the cancer comprises a myeloproliferative neoplasm. Accordingly, the polynucleotides, polypeptides and dimeric proteins may be formulated in a pharmaceutical composition as described herein. Such pharmaceutical composition may be used, *e.g.*, as a vaccine, *e.g.*, a cancer vaccine.

[00227] In some aspects, by the methods described herein it is possible to treat a patient suffering from cancer by examining any mutations present in the tumor in the patient, producing the pharmaceutical composition, *e.g.*, vaccine, *e.g.*, vaccine comprising a calreticulin Vaccibody™ as disclosed herein, and then immunizing the patient with the vaccine directed exactly to neoantigens present in his or her tumor, *e.g.*, directed against calreticulin frameshift neoantigens. The cancer may be a primary tumor, metastasis or both. The tumor examined for mutations may be a primary tumor or a metastasis. In some embodiments, the treatment comprises administering a pharmaceutical composition comprising a polynucleotide as described herein.

[00228] In some aspects, the patient is treated with a pharmaceutical composition, *e.g.*, vaccine, comprising a combination of peptides, *e.g.*, directed against calreticulin frameshift neoantigens.

[00229] In some embodiments, when the pharmaceutical composition comprises a polynucleotide of the disclosure, the pharmaceutical composition is administered intramuscularly.

[00230] Any suitable method for injecting the polynucleotide may be used, such as by the use of a jet injector or assisted by electroporation.

[00231] A pharmaceutical composition of the disclosure be administered as a single dosage, or may be administered in repeated doses. When administration is repeated, subsequent doses may be administered with at least 3-week intervals between dosing, to avoid exhaustion of the T cells.

[00232] Accordingly, in one embodiment the dosage regimen would be vaccination week 0, 3, 6 and then every 4 weeks as long as the patient has clinical benefit. The pharmaceutical composition may be administered for at least a year.

[00233] The pharmaceutical composition is administered in an immunologically effective amount. By "immunologically effective amount" is meant the amount required to establish a tumor reducing effect. Ultimately, the physician determines the dosage that typically is in the range of 0.3-6 mg for DNA-based compositions, and in the range of 5 μg-5 mg for polypeptide/protein compositions.

[00234] In another aspect, provided herein is a method of treating a disease or disorder associated with calreticulin frameshift expression, *e.g.*, a cancer.

[00235] In some embodiments, the cancer is a myeloproliferative neoplasm. In some embodiments, the myeloproliferative neoplasm is chronic myeloid leukemia, acute myeloid leukemia, polycythaemia vera, essential thrombocythemia, primary myelofibrosis, chronic eosinophilic leukemia, chronic myelomonocytic leukemia, systemic mastocytosis, idiopathic myelofibrosis, or myeloma.

[00236] In some embodiments, the cancer is a hematologic malignancy. In some embodiments, the hematologic malignancy is a leukemia. For example and without limitation, the leukemia may be acute myeloid leukemia (AML), acute lymphoblastic leukemia (ALL), myelodysplasia, myelodysplastic syndromes, acute T-lymphoblastic leukemia, or acute promyelocytic leukemia, chronic myelomonocytic leukemia, or myeloid blast crisis of chronic myeloid leukemia. In some embodiments, the cancer is a myeloproliferative neoplasm (MPN). In some embodiments, the MPN is chronic myeloid leukemia, acute myeloid leukemia, polycythaemia vera, essential thrombocythemia, primary myelofibrosis, chronic eosinophilic leukemia, chronic myelomonocytic leukemia, systemic mastocytosis, idiopathic myelofibrosis, or myeloma.

[00237] Any cancer expressing calreticulin, *e.g.*, a calreticulin frameshift mutation in exon 9, is envisaged as within the scope of the instant disclosure. In some embodiments, the cancer is a solid tumor. For example and without limitation, the cancer may be a brain cancer, bladder cancer, breast cancer, cervical cancer, colon cancer, colorectal cancer, endometrial

cancer, esophageal cancer, leukemia, lung cancer, liver cancer, melanoma, ovarian cancer, pancreatic cancer, prostate cancer, rectal cancer, renal cancer, stomach cancer, testicular cancer, or uterine cancer. Alternatively, the cancer may be a vascularized tumor, squamous cell carcinoma, adenocarcinoma, small cell carcinoma, melanoma, glioma, neuroblastoma, sarcoma *e.g.*, an angiosarcoma or chondrosarcoma), larynx cancer, parotid cancer, biliary tract cancer, thyroid cancer, acral lentiginous melanoma, actinic keratoses, acute lymphocytic leukemia, acute myeloid leukemia, adenoid cystic carcinoma, adenomas, adenosarcoma, adenosquamous carcinoma, anal canal cancer, anal cancer, anorectum cancer, astrocytic tumor, Bartholin gland carcinoma, basal cell carcinoma, biliary cancer, bone cancer, bone marrow cancer, bronchial cancer, bronchial gland carcinoma, carcinoid, cholangiocarcinoma, chondrosarcoma, choroid plexus papilloma/carcinoma, chronic lymphocytic leukemia, chronic myeloid leukemia, clear cell carcinoma, connective tissue cancer, cystadenoma, digestive system cancer, duodenum cancer, endocrine system cancer, endodermal sinus tumor, endometrial hyperplasia, endometrial stromal sarcoma, endometrioid adenocarcinoma, endothelial cell cancer, ependymal cancer, epithelial cell cancer, Ewing's sarcoma, eye and orbit cancer, female genital cancer, focal nodular hyperplasia, gallbladder cancer, gastric antrum cancer, gastric fundus cancer, gastrinoma, glioblastoma, glucagonoma, heart cancer, hemangioblastomas, hemangioendothelioma, hemangiomas, hepatic adenoma, hepatic adenomatosis, hepatobiliary cancer, hepatocellular carcinoma, Hodgkin's disease, ileum cancer, insulinoma, intraepithelial neoplasia, intraepithelial squamous cell neoplasia, intrahepatic bile duct cancer, invasive squamous cell carcinoma, jejunum cancer, joint cancer, Kaposi's sarcoma, pelvic cancer, large cell carcinoma, large intestine cancer, leiomyosarcoma, lentigo maligna melanomas, lymphoma, male genital cancer, malignant melanoma, malignant mesothelial tumors, medulloblastoma, medulloepithelioma, meningeal cancer, mesothelial cancer, metastatic carcinoma, mouth cancer, mucoepidermoid carcinoma, multiple myeloma, muscle cancer, nasal tract cancer, nervous system cancer, neuroepithelial adenocarcinoma nodular melanoma, non-epithelial skin cancer, non-Hodgkin's lymphoma, oat cell carcinoma, oligodendroglial cancer, oral cavity cancer, osteosarcoma, papillary serous adenocarcinoma, penile cancer, pharynx cancer, pituitary tumors, plasmacytoma, pseudosarcoma, pulmonary blastoma, rectal cancer, renal cell carcinoma, respiratory system cancer, retinoblastoma, rhabdomyosarcoma, sarcoma, serous carcinoma, sinus cancer, skin cancer, small cell carcinoma, small intestine cancer, smooth muscle cancer, soft tissue cancer, somatostatin-secreting tumor, spine cancer, squamous cell carcinoma, striated muscle cancer,

submesothelial cancer, superficial spreading melanoma, T cell leukemia, tongue cancer, undifferentiated carcinoma, ureter cancer, urethra cancer, urinary bladder cancer, urinary system cancer, uterine cervix cancer, uterine corpus cancer, uveal melanoma, vaginal cancer, verrucous carcinoma, VIPoma, vulva cancer, well differentiated carcinoma, or Wilms tumor.

Combination therapies

[00238] The disclosure provides for combination therapies. Compositions of the present disclosure may be combined with any other suitable treatment for cancer. In some embodiments, when the patient is administered a composition of the present disclosure for treatment of a cancer, the patient is further administered a radiation therapy, chemotherapy, surgical treatment, or combination thereof.

[00239] Example therapeutic agents that may be used as part of a combination therapy in treating cancer, include, for example and without limitation, radiation, mitomycin, tretinoin, ribomustin, gemcitabine, vincristine, etoposide, cladribine, mitobronitol, methotrexate, doxorubicin, carboquone, pentostatin, nitracrine, zinostatin, cetrorelix, letrozole, raltitrexed, daunorubicin, fadrozole, fotemustine, thymalfasin, sobuzoxane, nedaplatin, cytarabine, bicalutamide, vinorelbine, vesnarinone, aminoglutethimide, amsacrine, proglumide, elliptinium acetate, ketanserin, doxifluridine, etretinate, isotretinoin, streptozocin, nimustine, vindesine, flutamide, drogenil, butocin, carmofur, razoxane, sizofilan, carboplatin, mitolactol, tegafur, ifosfamide, prednimustine, picibanil, levamisole, teniposide, improsulfan, enocitabine, lisuride, oxymetholone, tamoxifen, progesterone, mepitiostane, epitiostanol, formestane, interferon-alpha, interferon-2 alpha, interferon-beta, interferon-gamma (IFN-γ), colony stimulating factor-1, colony stimulating factor-2, denileukin diftitox, interleukin-2, luteinizing hormone releasing factor and variations of the aforementioned agents that may exhibit differential binding to its cognate receptor, and increased or decreased serum half-life.

[00240] An additional class of agents that may be used as part of a combination therapy in treating cancer is immune checkpoint inhibitors. Example immune checkpoint inhibitors include, for example and without limitation, agents, such as antibodies, that inhibit one or more of (i) cytotoxic T lymphocyte-associated antigen 4 (CTLA4), (ii) programmed cell death protein 1 (PD l), (iii) PDLl , (iv) LAG3, (v) B7-H3, (vi) B7-H4, and (vii) TIM3. Example PD1 and PD-L1 inhibitors include, but are not limited to, Nivolumab, Pembrolizumab, Atexoliumab, Avelumab, Durvalumab, Cemiplimab, Dostarlimab and Retifanlimab. Example CTLA4 inhibitors include Ipilimumab, Botensilimab and

Tremelimumab. Example LAG3 inhibitors include Relatlimab. Example TIM3 inhibitors include Cobolimab.

[00241] Yet other agents that may be used as part of a combination therapy in treating cancer are monoclonal antibody agents that target non-checkpoint targets (*e.g.*, herceptin) and non-cytotoxic agents (*e.g.*, tyrosine-kinase inhibitors).

[00242] Yet other categories of anti-cancer agents include, for example and without limitation: (i) an inhibitor selected from an ALK Inhibitor, an ATR Inhibitor, an A2A Antagonist, a Base Excision Repair Inhibitor, a Bcr-Abl Tyrosine Kinase Inhibitor, a Bruton's Tyrosine Kinase Inhibitor, a CDC7 Inhibitor, a CHK1 Inhibitor, a Cyclin-Dependent Kinase Inhibitor, a DNA-PK Inhibitor, an Inhibitor of both DNA-PK and mTOR, a DNMT1 Inhibitor, a DNMTl Inhibitor plus 2-chloro-deoxyadenosine, an HDAC Inhibitor, a Hedgehog Signaling Pathway Inhibitor, an IDO Inhibitor, a JAK Inhibitor, a mTOR Inhibitor, a MEK Inhibitor, a MELK Inhibitor, a MTH1 Inhibitor, a PARP Inhibitor, a Phosphoinositide 3-Kinase Inhibitor, an Inhibitor of both PARP1 and DHODH, a Proteasome Inhibitor, a Topoisomerase-II Inhibitor, a Tyrosine Kinase Inhibitor, a VEGFR Inhibitor, and a WEE1 Inhibitor; (ii) an agonist of OX40, CD 137, CD40, GITR, CD27, HVEM, TNFRSF25, or ICOS; and (iii) a cytokine selected from IL-12, IL-15, GM-CSF, and G-CSF.

[00243] The person of skill in the art will appreciate that combination therapies are designed based upon the cancer to be treated, *e.g.*, MPN. MPN-specific therapies may include, for example and without limitation, JAK inhibitors (*e.g.*, ruxolitinib, fedratinib), cytoreductive agents (*e.g.*, hydroxyurea, interferons), aspirin, phlebotomy, medications and red blood cell transfusions for anemia, and stem cell transplants.

Kits

[00244] The present disclosure also provides a kit comprising a polynucleotide described herein, or protein or polypeptide encoded therein. In some embodiments, a kit comprises a polynucleotide described herein, or protein or polypeptide encoded therein, and instructions for the use thereof.

[00245] In some embodiments the kit further comprises a set of instructions comprising at least one method for treating a disease, *e.g.*, a cancer, *e.g.*, MPN, with the composition.

[00246] In some embodiments, the kits can optionally include a delivery vehicle for the composition, such as a syringe or injector. The reagents may be provided suspended in the excipient and/or delivery vehicle, or may be provided as a separate component which can be later combined with the excipient and/or delivery vehicle.

[00247] In some embodiments, the kits may optionally contain additional therapeutics to be co-administered. Suitable additional therapeutics can include those described in the combination therapies provided herein.

[00248] While the instructional materials typically comprise written or printed materials, they are not limited to such. Any medium capable of storing such instructions and communicating them to an end user is contemplated. Such media include but are not limited to electronic storage media (*e.g.*, magnetic discs, tapes, cartridges, chips), optical media (*e.g.*, CD ROM), and the like. Such media may include addresses to internet sites that provide such instructional materials.

[00249] In some embodiments, a kit comprises a unit dosage of a polynucleotide described herein, or dimeric protein or polypeptide encoded therein.

EXAMPLES

[00250] The disclosure now being generally described, will be more readily understood by reference to the following examples, which are included merely for purposes of illustration of certain aspects and embodiments of the present disclosure, and is not intended to limit the disclosure.

Example 1. Identification of potential CD8 T-cell epitopes for the CALR frameshift mutation

Methods

Sample Preparation

[00251] Samples of cultured HEK293 $CALR^{MUT}$ cells were lysed with 4 ml of lysis agent, consisting of 10% nonidet P-40 (NP-40, Cat no: S-004-C G BioScience®), 5M sodium chloride (Sigma®), 1M Tris hydrochloride with a pH of 8.0 (Cat no: T1080 Teknova®), 0.5M ethylenediamine tetra acetic acid (Cat no: E0306 Teknova®), protease inhibitor (Cat no: 78442 Thermo®), and water, which was added to the samples in the centrifuge tubes based on their weights. The samples were rotated at 4 °C for one hour and then stored at -20 °C overnight.

[00252] The cell lysate was thawed on ice and transferred into 5 ml centrifuge tubes to be centrifuged at 16,000 rotations per minute (RPM) for 30 minutes at 4 °C. A Bradford Assay (Cat no: 5000001 Bio-Rad®) was completed for each of the samples to determine protein concentration.

Immunoprecipitation

[00253] HLA-Pan Class I (W6/32) columns were prepared using NHS activated Sepharose 4 beads (GE Healthcare® 17090601) and a coupling buffer of 0.2M sodium bicarbonate and 0.5M sodium chloride. The columns were washed with 0.1M Tris hydrochloride with a pH of 8.5, and 0.1M acetate buffer. Affinity purification was performed under gravity and the flow-through was captured for further analysis (**FIGs. 1A-1B**). 0.1M glycine (Sigma®) pH 2.7 was used to elute bound HLA molecules under gravity (**FIGs. 1A-1B**). 0.1% trifluoroacetic acid (Cat no: LC485-1 Honeywell) was added to the glycine elute. The HLA-associated-peptides were eluted using Sep-Pak (Cat no: WAT054960 Waters®) with a two-step elution. HLA-specific peptides were eluted using 30% acetonitrile (Cat no: LC34967 Honeywell®)/0.1% trifluoroacetic acid and the HLA molecules were eluted using 70% acetonitrile/0.1% trifluoroacetic acid (**FIGs. 1A-1B**). Aliquots of the lysate, flow-through, glycine, 30% acetonitrile/0.1% trifluoroacetic acid, and 70% acetonitrile/0.1% trifluoroacetic acid eluates were collected throughout the process. Western analyses were performed on a few samples of interest.

[00254] Peptides and HLA fractions were placed on SpeedVac® Vacuum Concentrator (Thermo®) for 2 hours. Each sample, after SpeedVac®, was resuspended in 0.1% trifluoroacetic acid. Corresponding lung heavy peptides of interest were added to the samples in addition to 25 µl of 50 mM ammonium bicarbonate (Sigma®). The peptide fractions were purified further using a C-18 ZipTip® (Cat no: ZTC185096 Millipore®). 50 fmol of stable isotope labeled heavy synthetic analog of SPARPRTSC (SEQ ID NO:5) was added to the final peptide fraction. Stable isotope heavy synthetic analog has same physical properties so it will elute at the same retention time as endogenous peptide. All samples were then analyzed with the Orbitrap Fusion™ Lumos™ Tribrid™ Mass Spectrometer (Thermo®) for peptide sequencing.

LC-MS/MS analysis of HLA Class I peptides

[00255] HLA peptides as above were loaded onto a nanoViper® Acclaim PepMap100® C18 trap column (75 µm i.d. × 2 cm, 3 µm, 100 Å, Thermo®) and were separated using a nanoViper® Acclaim PepMap® RSLC C18 column (75 µm i.d. × 25 cm, 2 µm, 100 Å, Thermo®) heated to 40 °C and retrofitted with a New Objective SilicaTip® (7 cm) with a distal conductive coating at the inlet end of the emitter. The gradient was delivered by an EASY-nLC® 1200 HPLC system (Thermo®) at 300 nL/minute (min). The following 120-minute elution gradient with mobile phase A (Water/ 0.1% formic acid) and B (80%

Acetonitrile/ 0.1% formic acid) was used: 3% B at 3 min, linear to 35% B at 100 min, and linear to 45% B at 123 min. The peptides eluted from the column were ionized via Flex ion source at 1.9 kV and analyzed by the Thermo® Fusion Lumos Tribrid® mass spectrometer (Thermo®) using Xcalibur® 4.1.31.9 (Thermo®). Data acquisition was performed in data-dependent mode, where survey scans were carried out in the high field Orbitrap® analyzer (range of m/z 300-1500 at a resolution of 60,000) with the automatic gain control target of 4.0E5 and maximal ion fill time of 100 ms. The MS/MS analyses were performed by 1.2 m/z precursor ion isolation with the quadrupole, applying normalized HCD (higher-energy collisional dissociation) collision energy of 32%, and analysis of fragment ions in the Orbitrap® at a resolution of 15,000. Dynamic exclusion window was set to 6 seconds, monoisotopic precursor selection (MIPS) to peptide, maximum injection time to 100 milliseconds (ms), and charge states unknown. +1-+4 charge states were included, and the advanced peak determination was toggled on.

[00256] For FAIMS-enabled experiments, the settings were identical except the FAIMS device was placed between the nanoelectrospray source and the mass spectrometer. FAIMS separations were performed with the following settings: inner and outer electrode temperature set to 100 °C (except where noted), FAIMS carrier gas flow of 5.0 L/min, asymmetric waveform with DV −5000 V, entrance plate voltage 250 V, and CV settling time of 25 ms. The FAIMS carrier gas is N2, and the ion separation gap is 1.5 mm. The noted CVs were applied to the FAIMS electrodes. For external stepping or single CV experiments, the selected CV was applied to all scans throughout the analysis. For internal CV stepping experiments, each of the selected CVs was applied to sequential survey scans and MS/MS cycles (1 second); the MS/MS CV was always paired with the appropriate CV from the corresponding survey scan.

Data Analysis

[00257] Raw data files from the Orbitrap Fusion™ Lumos™ (Thermo) LC/MS were searched with the PEAKS® Studio X (BSI) proteomics software against Human Uniprot Database as well as custom databases for proteins of interest. The PEAKS® Studio X (BSI) searches were exported and run through custom R-scripts to compile those peptides into segmented folders. Those exported folders possessed Excel files containing peptide counts, representative peptide distributions and lists of all detected peptides with their abundances, ion scores, amino acid lengths, and masses.

[00258] The enrichment of HLA-class I was conducted on HEK293 CALRMUT cell lines using the W6/32 HLA-Class I pan antibody. The enriched HLA-bound peptides were subsequently analyzed using an advanced Orbitrap Fusion Lumos® mass spectrometer, which was coupled to a nanoLC system. The raw mass spectrometry (MS) data files were then searched with the UniProtKB (Uniprot knowledgebase release 2020_0420).

[00259] As shown in **FIG. 1A**, from each sample, approximately 8,000 peptides were identified, with an average detection of 4,500 9mers per sample. The distribution of peptide length was found to be consistent across all samples, with over 50% of the peptides being 9 amino acids in length. This is in line with expectations for HLA class I associated peptides.

[00260] Notably, SPARPRTSC (SEQ ID NO:5) peptides from calreticulin (CALR) L367fs*46, a CALR mutant, were detected in the engineered HEK293 $CALR^{MUT}$ cell line, as shown in **FIG. 1B**. The estimated copy number was determined to be approximately 550 copies per cell by comparing the MS1 intensity of mass spectra of the endogenous peptides with that of the stable isotope labeled (heavy Spike-In) peptides.

[00261] Accordingly, IP-MS identified an HLA-B*07:02 restricted epitope in engineered cancer cells, indicating $CALR^{MUT}$ is processed and presented on HLA.

Example 2. $CALR^{MUT}$-specific T cells were detected in the peripheral blood of myeloproliferative neoplasm (MPN) patients harboring the CALR exon 9+1 frameshift mutation

[00262] PBMCs isolated from patients were cultured in media containing IL-4, GM-CSF, and IL-21. IFNα was added following peptide pulsing. IL-7, IL-15, and IL-2 were added to the culture to support memory T cell expansion. Culture medium containing cytokines was replenished every 2 to 3 days until T cell re-stimulation. A schematic depiction of the experimental set-up is shown in **FIG. 2A**.

[00263] PBMCs from MPN patients carrying $CALR^{MUT}$ or a healthy donor were cultured in the presence of a peptide pool derived from $CALR^{MUT}$ or DMSO. Peptide pools are provided in **Tables 3** and **4** for calreticulin frameshift mutation and wildtype peptides, respectively. For Table 3, amino acids in italic are found in wild type CALR. Amino acids indicated in bold are the core sequence (36 amino acids in total) found in all MPN patients carrying CALR exon9 frameshift mutants. Expanded PBMC cultures were re-stimulated by autologous B cells pulsed with the $CALR^{MUT}$ peptide pool used for T cell expansion. For negative controls, expanded PBMC cultures were re-stimulated with B cells treated with DMSO or a peptide pool derived from wild-type CALR ($CALR^{WT}$). For a positive control,

expanded PBMC cultures were re-stimulated with B cells pulsed with a peptide pool derived from common viral proteins (CEF: CMV, EBV, and Flu). The CEF peptide pool details are provided in **Table 5** below.

[00264] $CD8^+$ T cell activation was measured by ICS and fluorescence activated cell sorting (FACS) for IFNγ and TNFα production. Representative intracellular cytokine staining (ICS) analysis is shown in **FIG. 2B**, and summarized data is shown in **FIG. 2C**, for IFNγ and TNFα produced by pre-expanded T cells upon antigen re-stimulation.

[00265] Taken together, these data demonstrate pre-existing $CALR^{MUT}$-specific T cells were detected in the peripheral blood of MPN patients.

Table 3. Amino acid sequences of calreticulin frameshift mutation peptides.

Peptide	$CALR^{MUT}$ Amino acid sequences	SEQ ID NO
CALR exon9 +1 frameshift resulting 44mer sequence	RRMMRTKM**RMRRMRRTRRKMRRKMSPARPRTSCREACLQGWTEA**	3
CALR_MUT-p1	*MKDKQDEEQR*TRRMM	12
CALR_MUT-p2	*DEEQR*TRRMMRTKM**R**	13
CALR_MUT-p3	TRRMMRTKM**RMRRMR**	14
CALR_MUT-p4	RTKM**RMRRMRRTRRK**	15
CALR_MUT-p5	**MRRMRRTRRKMRRKM**	16
CALR_MUT-p6	**RTRRKMRRKMSPARP**	17
CALR_MUT-p7	**MRRKMSPARPRTSCR**	18
CALR_MUT-p8	**SPARPRTSCREACLQ**	19
CALR_MUT-p9	**RTSCREACLQGWTEA**	20
CALR_MUT-p10v2	RKEEEEAEDNCRRMM	21
CALR_MUT-p11v2	EAEDNCRRMMRTKMR	22

Table 4. Amino acid sequences of calreticulin wildtype peptides.

Peptide	$CALR^{WT}$ Amino acid sequences	SEQ ID NO
CALR WT reference sequence	MKDKQDEEQRLKEEEEDKKRKEEEEAEDKEDDEDKDEDEEDEEDK EEDEEEDVPGQAKDEL	23
CALR_WT_p1	MKDKQDEEQRLKEEE	24
CALR_WT_p2	DEEQRLKEEEEDKKR	25
CALR_WT_p3	LKEEEEDKKRKEEEE	26
CALR_WT_p4	EDKKRKEEEEAEDKE	27
CALR_WT_p5	KEEEEAEDKEDDEDK	28
CALR_WT_p6	AEDKEDDEDKDEDEE	29
CALR_WT_p7	DDEDKDEDEEDEEDK	30

CALR_WT_p8	DEDEEDEEDKEEDEE	31
CALR_WT_p9	DEEDKEEDEEEDVPG	32
CALR_WT_p10	KEEDEEEDVPGQAK	33
CALR_WT_p11	DEEEDVPGQAKDEL	34

Table 5. Amino acid sequences of CEF peptides.

HLA Allele	Virus	Peptide Sequence	SEQ ID NO
A1	Influenza A	VSDGGPNLY	35
	Influenza A	CTELKLSDY	36
A2	Influenza M	GILGFVFTL	37
	Influenza A	FMYSDFHFI	38
	EBV LMP2A	CLGGLLTMV	39
	EBV BMLF1	GLCTLVAML	40
A0201	HCMV pp65	NLVPMVATV	41
AA68	Influenza NP	KTGGPIYKR	42
A3	Influenza NP	RVLSFIKGTK	43
	Influenza A	ILRGSVAHK	44
	EBV	RVRAYTYSK	45
	EBV	RLRAEAQVK	46
A3, A11, A60B1	Influenza M	SIIPSGPLK	47
A11	EBV EBNA 4NP	AVFDRKSDAK	48
	EBV	IVTDFSVIK	49
	EBV	ATIGTAMYK	50
A24	EBV RTA	DYCNVLNKEF	51
B7	Influenza NP	LPFDKTTVM	52
	EBV	RPPIFIRRL	53
B8	Influenza NP	ELRSRYWAI	54
	EBV BZLF-1	RAKFKQLL	55
	EBV EBNA 3A	FLRGRAYGL	56
	EBV EBNA 3	QAKW RLQTL	57
B18	HCMV	SDEEEAIVAYTL	58
B27	Influenza NP	SRYW AIRTR	59
	Influenza M	ASCMGLIY	60
	EBV EBNA 3C	RRIYDLIEL	61
B35	EBV EBNA3A	YPLHEQHGM	62
	CMV pp65	IPSINVHHY	63
B44	EBV	EENLLDFVRF	64
	HCMV	EFFW DANDIY	65
B0702	HCMV	TPRVTGGGAM	66

Example 3. MPN antigen-specific T cells were expanded from healthy donor naïve T cells and were detectable in the human T Cell repertoire.

[00266] Monocyte-derived dendritic cells (moDC) were derived from CD14+ cells isolated from healthy donor PBMCs in the presence of IL-4 and GM-CSF. TIIP cocktail (*i.e.*, TNFα, IL-1, IL-6, and PGE1) was added on day -1 following peptide pulsing of moDC on

day -2. Autologous naïve T cells isolated from PBMCs were added to the culture on day 0, in the presence of IL-21. IL-7 and IL-15 were added to the culture on day 3. Media containing cytokines was replenished every 2 to 3 days. The priming-expansion procedure was repeated once to expand low frequency precursor cells. A schematic depiction of the experimental set-up is shown in **FIG. 3A**.

[00267] Naïve T cells from healthy donors were expanded with a peptide pool derived from $CALR^{MUT}$ as described *supra*. Naïve T cells co-cultured with moDC-treated with DMSO were cultured in parallel as a negative control. Pre-expanded T cells were restimulated by CD2-depleted autologous PBMCs pulsed with the indicated peptide pools or DMSO. IFNγ-producing T cells were detected by IFNγ ELISpot. Representative ELISpot analysis is shown in **FIG. 3B**, and summarized data is shown in **FIG. 3C**, for $CALR^{MUT}$-specific T cells expanded from healthy donor naïve T cells.

[00268] Accordingly, $CALR^{MUT}$-specific T cell expansion can be induced from naïve T cells.

Example 4. Vaccibodies™ encoding MPN neoantigens elicited CALR-specific T cell responses in C57BL/6 mice.

[00269] C57BL/6 mice were implanted with $2x10^5$ engineered TC-1 cells on Day 0, and immunized either with CALR-44mer Vaccibody™, empty Vaccibody™ as a negative control, or HPV16 pDNA as a positive control. The CALR-44mer Vaccibody™ (CV002) consists of a targeting unit (MIP-1α chemokine), a dimerization unit, and the antigenic unit encoding the 44-amino acid sequence resulting from the CALR L367fs*46 mutation that is common among 80% of patients (*see* **Table 6** below, SEQ ID NO: 3). The empty Vaccibody™ has the same backbone but lacks the antigen sequence. The HPV16 E6 E7 pDNA is a positive control encoding the full protein. 20 µg or 60 µg of CALR-44mer Vaccibody™ or appropriate controls were administered intramuscularly followed by electroporation around the injection site at day 7 and boosted at day 14 post tumor implantation. At day 21, the spleens were harvested and processed into single cell suspensions, restimulated with a 15-mer peptide pool or DMSO for 24 h, and assessed for antigen specific T cell responses by IFNγ enzyme linked immune absorbent spots (ELISpots), shown in **FIG. 4**. For the ELISpot, isolated splenocytes were plated at $3x10^5$ cells per well and stimulated overnight with a pool of 15mer peptides covering the entirety of the CALR44mer neoantigen. Cells were discarded the following day and the IFNγ ELISpot was performed following manufacturer protocols (Immunospot).

[00270] Restimulation peptides sequences used are provided in **Table 6**.

[00271] As shown in FIG. 4, immunization with Vaccibody™ pDNA induced significant cellular immune responses in tumor-bearing animals vaccinated with either 20 μg or 60 μg, with the magnitude of the response varying among individual animals. No immune responses were observed in non-tumor-bearing mice. These data demonstrate that immunization with CALR44mer Vaccibody™ elicits strong antigen-specific T cell responses in tumor bearing mice.

Example 5. An engineered mouse tumor cell line, TC-1, and was functionally regressed with therapeutic MPN Vaccibody™ immunization.

[00272] To test the efficacy of the CALR 44-mer Vaccibody™, a tumor model was established using an engineered TC-1 (a C57BL/6 mouse lung epithelial cell line) expressing the CALR 44-mer amino acid sequence described in the above Examples (SEQ ID NO:3). This neoantigen is present in $CALR^{MUT}$ patients carrying insertion/deletion (indels) mutations in exon 9, which result in a frameshift and the generation of a new C-terminal of the calreticulin protein containing the conserved 44mer neoantigen. The expression construct used to engineer this cell line included an HA-tagged peptide ubiquitin cassette for CALR frameshift mutation and Blasticidin resistance gene. The cells were transduced with a pLVX lentivirus using Lenti-X packaging single shot (VSV-G; Clontech™, #631276). The transduced cells were positively selected *via* Blasticidin selection, single cell sorted, and expanded. The expanded single clone populations were validated by intracellular staining with Alexa Fluor 647 conjugated HA-tag antibody (R&D Systems™, Cat# IC6875R) using flow cytometry, as shown in **FIG. 5A.**

[00273] A total of about $2.5x10^5$ engineered TC-1 cells were implanted subcutaneously on the right flank of each mouse. Five mice were used for each group. Seven days post-tumor implantation, 2, 20 or 60 μg of pDNA vaccines (CALR-44mer Vaccibody™ carrying the CALR44mer neoantigen (SEQ ID NO: 3) or empty Vaccibody™ without the 44mer neoantigen) were injected intramuscularly and electroporated into the mouse right upper leg. The immunization was followed by a booster dose on day 14. Tumor growth was monitored twice weekly over a 40-day period.

[00274] As shown in **FIG. 5B**, following the initial growth of TC-1 tumors to approximately 350-500 mm^3 by day 22, mice that received two vaccinations with CALR44mer Vaccibody™ pDNA demonstrated significant tumor regression in a dose-dependent manner. The most pronounced tumor regression was observed in animals vaccinated with a 20 μg dose, which resulted in an approximately 65% reduction in tumor

size in average post-vaccine boost. Immunization with CALR-44mer Vaccibody™ elicits potent anti-tumor protection in a dose response manner in mice transplanted with TC-1 engineered cells.

[00275] Taken together, these data demonstrate that immunization with the CALR 44-mer Vaccibody™ generated CALR-specific T cell responses as measured by IFNγ production and functionally regressed an engineered mouse tumor cell line, TC-1 in C57BL/6 mice.

Table 6. Example calreticulin frameshift mutation peptides.

Peptide	$CALR^{MUT}$ Amino acid sequences	SEQ ID NO
CALRfs resulting 44mer sequence	RRMMRTKM**RMRRMRRTRRKMRRKMSPARPRTSCREACLQGWTEA**	3
15mer pool ovlp-1	RRMMRTKM**RMRRMRR**	67
15mer pool ovlp-2	RTKM**RMRRMRRTRRK**	15
15mer pool ovlp-3	**RMRRMRRTRRKMRRK**	68
15mer pool ovlp-4	**MRRTRRKMRRKMSPA**	69
15mer pool ovlp-5	**RRKMRRKMSPARPRT**	70
15mer pool ovlp-6	**RRKMSPARPRTSCRE**	71
15mer pool ovlp-7	**SPARPRTSCREACLQ**	19
15mer pool ovlp-8	**PRTSCREACLQGWTEA**	72

EQUIVALENTS

[00276] The details of one or more embodiments of the disclosure are set forth in the accompanying description above. Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present disclosure, the preferred methods and materials are now described. Other features, objects, and advantages of the disclosure will be apparent from the description and from the claims. In the specification and the appended claims, the singular forms include plural referents unless the

context clearly dictates otherwise. Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this disclosure belongs. All patents and publications cited in this specification are incorporated by reference.

[00277] The foregoing description has been presented only for the purposes of illustration and is not intended to limit the disclosure to the precise form disclosed, but by the claims appended hereto.

WHAT IS CLAIMED IS:

1. A pharmaceutical composition comprising:

a polynucleotide comprising one or more sequences encoding a targeting unit, one or more sequences encoding a dimerization unit, and one or more sequences encoding an antigenic unit,

wherein the antigenic unit comprises a sequence encoding at least one epitope derived from calreticulin, and

a pharmaceutically acceptable carrier.

2. The pharmaceutical composition of claim 1, wherein the at least one epitope derived from calreticulin comprises a sequence of a frameshift mutation in calreticulin exon 9.

3. The pharmaceutical composition of claim 1 or 2, wherein the at least one epitope derived from calreticulin comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 4-11.

4. The pharmaceutical composition of claim 2 or 3, wherein the frameshift mutation in calreticulin comprises an amino acid sequence of SEQ ID NO: 2-3 or 141-158.

5. The pharmaceutical composition of any one of claims 1-4, wherein the at least one epitope has a length of 7 to 44 amino acids.

6. The pharmaceutical composition of any one of claims 1-5, wherein the antigenic unit comprises at least two, at least three, at least four, or at least five epitopes from calreticulin.

7. The pharmaceutical composition of any one of claims 1-6, wherein the dimerization unit comprises a hinge region.

8. The pharmaceutical composition of claim 7, wherein the hinge region is derived from an immunoglobulin.

9. The pharmaceutical composition of claim 7 or 8, wherein the dimerization unit further comprises a dimerization domain.

10. The pharmaceutical composition of claim 9, wherein the dimerization domain comprises an immunoglobulin constant domain.

11. The pharmaceutical composition of claim 9 or 10, wherein the immunoglobulin constant domain is a carboxyterminal C domain derived from an IgG.

12. The pharmaceutical composition of claim 11, wherein the carboxyterminal C domain is derived from IgG3.

13. The pharmaceutical composition of any one of claims 1-12, wherein the dimerization unit comprises a linker.

14. The pharmaceutical composition of claim 13, wherein the linker connects the hinge region and the dimerization domain.

15. The pharmaceutical composition of any one of claims 7-14, wherein the dimerization unit comprises hinge exons h1 and h4 connected through a linker to a CH3 domain of human IgG3.

16. The pharmaceutical composition of any one of claims 7-15, wherein the dimerization unit comprises a sequence having at least 80%, at least 90%, at least 95%, or at least 99% identity to SEQ ID NO:178.

17. The pharmaceutical composition of claim 16, wherein the dimerization unit comprises a sequence of SEQ ID NO:178.

18. The pharmaceutical composition of any one of claims 1-17, wherein the antigenic unit and the dimerization unit are connected by a first linker.

19. The pharmaceutical composition of claim 18, wherein the first linker comprises an amino acid sequence of SEQ ID NO:98.

20. The pharmaceutical composition of any one of claims 1-19, wherein the targeting unit comprises a moiety that interacts with a surface molecule on an antigen presenting cell.

21. The pharmaceutical composition of claim 20, wherein the surface molecule is selected from the group consisting of HLA, CD14, CD40, a chemokine receptor, and a Toll-like receptor.

22. The pharmaceutical composition of claim 20 or 21, wherein the targeting unit comprises a MIP-1α chemokine that binds to a CCR5 receptor.

23. The pharmaceutical composition of claim 22, wherein the MIP-1α chemokine comprises an amino acid sequence of SEQ ID NO:159.

24. The pharmaceutical composition of any one of claims 1-23, wherein the polynucleotide is an RNA or a DNA.

25. The pharmaceutical composition of claim 24, wherein the polynucleotide comprises a nucleotide sequence encoding a signal peptide.

26. The pharmaceutical composition of any one of claims 1-25, comprising, from 5′ to 3′, the targeting unit, the dimerization unit, and the antigenic unit.

27. The pharmaceutical composition of claim 26, wherein the targeting unit encodes a MIP-1α sequence, the dimerization unit encodes a hinge exon h1 from IgG3, a hinge exon h4 from IgG3, a dimerization unit linker, and human CH3 domain of IgG3, and the antigenic unit encodes an epitope derived from calreticulin comprising SEQ ID NOs:4-11.

28. The pharmaceutical composition of claim 27, wherein the polynucleotide comprises a sequence encoding a signal peptide of MIP-1α 5′ to the targeting unit, and a sequence encoding a linker between the dimerization unit and the antigenic unit.

29. The pharmaceutical composition of any one of claims 1-28, wherein the polynucleotide comprises, from 5′ to 3′:

a targeting unit encoding SEQ ID NO:159;

a dimerization unit encoding SEQ ID NO:178; and

an antigenic unit encoding SEQ ID NO:3.

30. The pharmaceutical composition of claim 29, wherein the polynucleotide comprises, from 5′ to 3′:

a signal peptide encoding SEQ ID NO:163;

a targeting unit encoding SEQ ID NO:159;

a dimerization unit encoding SEQ ID NO:178;

a linker encoding SEQ ID NO:98; and

an antigenic unit encoding SEQ ID NO:3.

31. A pharmaceutical composition comprising a polynucleotide encoding an amino acid sequence of SEQ ID NO:177.

32. A polynucleotide of any one of claims 1-31.

33. An expression vector comprising the polynucleotide of claim 32.

34. A host cell comprising the polynucleotide of claim 32 or the expression vector of claim 33.

35. A polypeptide encoded by the polynucleotide of claim 32.

36. A dimeric protein comprising two polypeptides of claim 35.

37. The dimeric protein of claim 36, wherein the dimeric protein is a homodimer.

38. The polynucleotide of claim 32, the polypeptide of claim 35, or the dimeric protein of claim 36 or 37 for use in the manufacture of a medicament for treating a cancer.

39. The pharmaceutical composition of any one of claims 1-31, for use in a method of treating cancer in a subject.

40. The pharmaceutical composition of any one of claims 1-31, for use in the manufacture of a medicament for treating cancer in a subject.

41. A method of treating cancer in a subject in need thereof, comprising administering a therapeutically effective amount of the pharmaceutical composition of any one of claims 1-31 to the subject.

42. A method of eliciting an immune response in a subject in need thereof, comprising administering the pharmaceutical composition of any one of claims 1-31 to the subject.

43. A method of treating cancer in a subject in need thereof, the method comprising administering to the subject a pharmaceutical composition comprising:

a polynucleotide comprising one or more sequences encoding a targeting unit, one or more sequences encoding a dimerization unit, and one or more sequences encoding an antigenic unit,

wherein the antigenic unit comprises a sequence encoding at least one epitope derived from calreticulin, and

a pharmaceutically acceptable carrier.

44. The pharmaceutical composition for use of any one of claims 38-40 or the method of any one of claims 41-43, wherein the cancer is a myeloproliferative neoplasm.

45. The pharmaceutical composition for use or the method of claim 44, wherein the myeloproliferative neoplasm is chronic myeloid leukemia, acute myeloid leukemia,

polycythaemia vera, essential thrombocythemia, primary myelofibrosis, chronic eosinophilic leukemia, chronic myelomonocytic leukemia, systemic mastocytosis, idiopathic myelofibrosis, or myeloma.

46. The pharmaceutical composition for use or the method of any one of claims 38-45, wherein the cancer expresses calreticulin.

47. The pharmaceutical composition for use or the method of claim 46, wherein the calreticulin comprises a frameshift mutation in exon 9.

48. The pharmaceutical composition for use or the method of any one of claims 38-47, wherein the treatment is therapeutic.

49. The pharmaceutical composition for use or the method of any one of claims 38-47, wherein the treatment is prophylactic.