



(51) International Patent Classification:

C07K 16/28 (2006.01) A61P 35/00 (2006.01)

C07K 16/30 (2006.01) C07K 14/435 (2006.01)

C12N 15/52 (2006.01) C07K 14/725 (2006.01)

C12N 15/62 (2006.01) C07K 14/705 (2006.01)

DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

(21) International Application Number:

PCT/GB2023/050997

Published:

— with international search report (Art. 21(3))

— with sequence listing part of description (Rule 5.2(a))

(22) International Filing Date:

13 April 2023 (13.04.2023)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

2205589.1 14 April 2022 (14.04.2022) GB

(71) Applicant: CRESCENDO BIOLOGICS LIMITED

[GB/GB]; Meditrina Building 260, Babraham Research Campus, Cambridge CB22 3AT (GB).

(72) Inventors: JOHNSTON, Colette; Crescendo Biologics

Limited, Meditrina Building 260, Babraham Research Campus, Cambridge CB22 3AT (GB). HAYES, Phil; Crescendo Biologics Limited, Meditrina Building 260, Babraham Research Campus, Cambridge CB22 3AT (GB).

(74) Agent: APPLEYARD LEES IP LLP; 15 Clare Road, Halifax

Yorkshire HX1 2HY (GB).

(81) Designated States (unless otherwise indicated, for every

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

(84) Designated States (unless otherwise indicated, for every

kind of regional protection available): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ,

(54) Title: MESOTHELIN BINDERS

Figure 1

A

J591 V_L V_H FH TM CD8 CD3ζPSMA-V_H V_H FH TM CD8 CD3ζ

(57) Abstract: The invention relates to isolated single domain antibodies that bind mesothelin, related multispecific molecules, pharmaceutical compositions and methods of treatment.



Mesothelin binders

Field of the invention

The invention relates to single domain antibodies that bind mesothelin and can be used in treating cancer.

Introduction

Mesothelin is a tumour differentiation antigen that is normally present on the mesothelial cells lining the pleura, peritoneum and pericardium. It is however highly expressed in several human cancers including malignant mesothelioma, pancreatic, ovarian, endometrial and lung adenocarcinoma. In the context of cancer, high expression levels of MSLN have been
10 correlated with poor prognosis in ovarian cancer, cholangiocarcinoma, lung adenocarcinoma and triple-negative breast cancer. The limited expression of mesothelin on normal human tissues and high expression in several human cancers makes mesothelin an attractive candidate for cancer therapy. Several therapeutic strategies have been designed for targeting mesothelin on tumor cells including tumor vaccine strategy, antibody-based therapies and adoptive CAR T-cell therapy. These therapies are being evaluated in phase I and/or phase II clinical trials.

Heavy-chain-only antibodies (HcAbs) without light chains have been reported in camelids and camelids and cartilaginous fish, and shown to exhibit strong and specific antigen binding (Holliger P and Hudson PJ. Engineered antibody fragments and the rise of single domains.
20 Nat Biotechnol 2005;23:1126–36).

Summary of the Invention

In a first aspect, the invention relates to an isolated V_H single domain antibody that specifically binds to human mesothelin (MSLN), wherein the V_H single domain antibody comprising a CDR1 comprising SEQ NO. 4 or a sequence with 1, 2 or 3 amino acid modifications, CDR2 comprising SEQ NO. 5 or a sequence with 1, 2, 3 or 4 amino acid modifications and a CDR3 comprising SEQ NO. 6 or a sequence with 1, 2 or 3 amino acid modifications.

In another aspect, the invention relates to an isolated nucleic acid encoding a V_H single domain antibody as described herein.

In another aspect, the invention relates to a vector comprising a nucleic acid as described
30 herein.

In another aspect, the invention relates to a host cell comprising a nucleic acid or a vector as described herein.

In another aspect, the invention relates to a binding molecule comprising an isolated V_H single domain antibody as described herein.

In another aspect, the invention relates to a use of a single domain antibody as described herein in a multispecific or multivalent binding agent.

In another aspect, the invention relates to a pharmaceutical composition comprising a V_H single domain antibody or a binding molecule as described herein and a pharmaceutical acceptable carrier, excipient or diluent.

- 10 In another aspect, the invention relates to a method for treating a cancer comprising administering a V_H single domain antibody or a binding molecule as described herein.

In another aspect, the invention relates to a V_H single domain antibody, a pharmaceutical composition or a binding molecule as described herein for use in the treatment of cancer.

In another aspect, the invention relates to an immunoconjugate comprising a single domain antibody according or a binding molecule as described herein linked to a therapeutic agent.

In another aspect, the invention relates to a method for producing a V_H single domain antibody as described herein comprising expressing a nucleic acid encoding said binding molecule in a host cell and isolating the binding molecule from the host cell.

- 20 In another aspect, the invention relates to a kit comprising a single domain antibody or a binding molecule, or a pharmaceutical composition as described herein and optionally instructions for use.

In another aspect, the invention relates to a method for detecting the presence of human MSLN in a test sample or detecting or diagnosing a cancer in a subject comprising contacting said sample with a single domain antibody as described herein and at least one detectable label and detecting binding of said single domain antibody to human MSLN.

In another aspect, the invention relates to a method for producing a V_H single domain antibody that binds to human MSLN comprising

- 30 a) immunising a transgenic animal that expresses a nucleic acid construct comprising human heavy chain V genes and that is not capable of making functional endogenous light or heavy chains with an MSLN antigen,

- b) generating a library from said animal
- c) isolating single V_H domain antibodies from said libraries and
- d) identifying a single V_H domain antibody that binds to human MSLN and
- e) isolating said antibody.

In another aspect, the invention relates to a single V_H domain antibody obtained or obtainable by the method above.

In another aspect, the invention relates to an isolated heavy chain only antibody comprising a V_H domain that binds to human MSLN.

10 In another aspect, the invention relates to a heavy chain only antibody comprising a V_H domain that binds to human MSLN obtained or obtainable from a transgenic mouse which expresses human V, D and J loci and does not produce functional endogenous lambda and kappa light chains and heavy chain.

In another aspect, the invention relates to an antibody drug conjugate comprising a V_H single domain antibody or a binding molecule as described herein.

Figures

The invention is further described in the following non-limiting figures.

20 **Figure 1. Human antibody V_H domain-based CAR targeting PSMA is expressed and signals in T cells.** (A) Schematic diagram of J591 and PSMA-V_H constructs. (B, C) Representative flow cytometry plots (B) and summary (C) illustrating J591 and PSMA-V_H expression in T cells. The CD19-specific CAR (CD19) and non-transduced T cells (NT) were used as positive and negative controls, respectively. ****p < 0.0001, One-way ANOVA. (D) *In vitro* expansion of CD19, J591, PSMA-V_H and NT T cells; error bars represent SD, (n = 4). p ≥ 0.05 by One-way ANOVA. (E) T cell subset composition based on CD45RA and CCR7 expression in CD19, J591, PSMA-V_H and NT T cells at day 14 of culture; error bars represent SD, (n = 4). p ≥ 0.05 by One-way ANOVA. (F) Western blots detecting phosphorylation of CAR-CD3ζ, Akt and ERK in J591 and PSMA-V_H T cells activated via CAR cross-linking with an anti-FLAG Ab followed by incubation with a secondary Ab to induce the aggregation of CAR molecules. Total CAR, CD3ζ and endogenous CD3ζ were used as loading controls. Data are representative of 2 experiments.

Figure 2. T cells expressing the human antibody V_H domain-based CAR targeting PSMA are functional *in vitro*. (A) Representative flow cytometry plots showing the expression of PSMA in C4-2, PC3 and PC3 cells engineered with a retroviral vector to express PSMA. (B, C) Representative flow cytometry plots (B) and summary (C) illustrating Granzyme-B expression of T cells expressing either J591 or PSMA-V_H cocultured overnight with a tumor cell line expressing PSMA (PC3-PSMA-eGFP) at E:T ratio of 1:2; error bars represent SD, (n = 4). $p \geq 0.05$ by t-test. (D, E) Representative flow cytometry plots (D) and summary (E) illustrating the kinetics of CD69 expression of T cells expressing either J591 or PSMA-V_H and cocultured overnight with a tumor cell line expressing PSMA (PC3-PSMA-eGFP) at E:T ratio of 1:2. Data are representative of 4 experiments. *** $p < 0.001$ Two-way ANOVA. (F) Representative flow cytometry plots showing coculture of CD19, J591 and PSMA-V_H T cells with C4-2-eGFP, PC3-eGFP and PC3-PSMA-eGFP. T cells were cocultured with tumor cells at an E:T ratio of 1:5 for 6 days. At day 6, all cells were collected and analyzed by flow cytometry to quantify tumor cells (GFP) and T cells (CD3), respectively. (G) Summary of coculture of CD19, J591 and PSMA-V_H T cells with tumor cells in (F); error bars represent SD, (n = 4). **** $p < 0.0001$, Two-way ANOVA. (H, I) IFN- γ (H) and IL-2 (I) were detected by ELISA in the coculture supernatant of cocultures of CD19, J591 and PSMA-V_H T cells with tumor cells illustrated in (F); error bars represent SD, (n = 5). * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$, Two-way ANOVA. (J) Representative flow cytometry plots showing the proliferation of J591 and PSMA-V_H T cells in response to tumor cells as assed by CFSE dilution. Data are representative of 4 experiments.

Figure 3. T cells expressing the human antibody V_H domain-based CAR targeting PSMA are functional *in vivo*. (A) Schematic of the metastatic prostate cancer model using PC3-PSMA-FFluc-eGFP tumor cells in NSG mice and treatment with CD19, J591 and PSMA-V_H T cells (n = 5 mice per group). (B) Representative images of tumor bioluminescence (BLI) at selected time points post T cell injections. (C) Kinetics of tumor BLI post T cell injections. (D) Schematic of the metastatic prostate cancer model using PC3-PSMA-FFluc-eGFP tumor cells in NSG mice and treatment with low dose of CD19, J591 and PSMA-V_H T cells (n ≥ 4 mice per group). (E) Representative images of tumor BLI at selected time points post T cell injections for low dose of T cells. (F) Kinetics of tumor BLI post T cell injections low dose of T cells. (G, H) Percentage of T cells in the gate of live cells (G) and total cell numbers (H) in blood, spleen, and bone marrow from PC3-PSMA-bearing mice treated with low doses of CAR-T cells. Mice were euthanized at day 58 after CAR-T cells infusion and T cells were identified as CD45⁺CD3⁺ cells by flow cytometry. J594 group (n = 5) PSMA-V_H group (n = 4). $p \geq 0.05$ by t-test.

Figure 4. T cells expressing the human antibody V_H domain-based CAR targeting MSLN demonstrate antitumor activity. (A) Schematic diagram of MSLN-scFv and MSLN-heavy-

chain-only (MSLN-V_H) CAR constructs. (B) Summary of coculture of CD19, MSLN.scFv and MSLN-V_H T cells with Aspc-1-eGFP (MSLN⁺) and PC3-eGFP (MSLN⁻) tumor cell lines. T cells were cocultured with tumor cells at an E:T ratio of 1:5 for 6 days. At day 6, all cells were collected and analyzed by flow cytometry to quantify tumor cells and T cells, respectively. Error bars represent SD, (n = 4). ****p < 0.0001, Two-way ANOVA. (C) IFN-γ (upper panel) and IL-2 (lower panel) detected in the supernatants of the cocultures illustrated in (B) as measured by ELISA; error bars represent SD, (n = 4). ****p < 0.0001, Two-way ANOVA. (D) Representative flow cytometry plots showing the proliferation of MSLN.scFv and MSLN-V_H T cells in response to tumor cells as assessed by CFSE dilution. Data are representative of 3 experiments. (E) Schematic of the metastatic pancreatic cancer model using Aspc-1-FFluc-eGFP tumor cells in NSG mice. (F, G) Representative tumor BLI (F) and bioluminescence kinetics (G) of Aspc-1-FFluc-eGFP tumor growth at the representative time points post T cell injections. (n = 5 mice per group). (H) Kaplan-Meier survival curve of mice in (E) (n = 5 mice per group). Data are representative of two experiments. (I) Frequency of human CD45⁺CD3⁺ cells in blood at 22 days (left) post T-cell infusion and at euthanasia (right) of MSLN-scFv and MSLN-V_H T cells, respectively. Data are shown as individual values and the mean (n = 5 mice per group). p ≥ 0.05 by t-test.

Figure 5. *In vitro* analysis of monospecific and bispecific Humabody V_H binding. (A) Schematic representation of monospecific (single V_H) or bispecific (double V_H) proteins. (B) Single cycle Biacore kinetic analysis of PSMA binding. (C) Biacore kinetic analysis of MSLN binding, 3-fold dilution series starting at 300nM, except the control scFv protein which started at 33.3nM. Data are representative of two experiments.

Figure 6. T cells expressing two human antibody V_H domain-based CARs demonstrate dual specificity *in vitro*. (A) Schematic diagram of PSMA-V_H, MSLN-V_H, and PSMA/MSLN-V_H CAR constructs. (B, C) Representative flow cytometry plots (B) and summary (C) illustrating CAR expression in T cells. The CD19-specific CAR (CD19) was used as negative controls. p ≥ 0.05 by One-way ANOVA. (D) Representative flow cytometry plots showing PC3-PSMA-eGFP (PSMA target), PC3-MSLN-eGFP (MSLN target), and mixture of PC3-PSMA-eGFP and PC3-MSLN-eGFP (1:1 ratio) cotultured with CD19.CAR, PSMA-V_H.CAR, MSLN-V_H.CAR and PSMA/MSLN-V_H.CAR T cells at the E: T ratio of 1:5 for 6 days. Tumor cells and T cells were quantified at day 6 by flow cytometry. (E) Summary of coculture experiments illustrated in (D); error bars represent SD, (n = 5). *p < 0.05, **p < 0.01, ****p < 0.0001, Two-way ANOVA. (F, G) IFN-γ (F) and IL-2 (G) detected in the coculture supernatant of the coculture experiments described in (D) as measured by ELISA; error bars represent SD, (n = 3) *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001, Two-way ANOVA.

Figure 7. T cells expressing two human antibody V_H domain-based CARs demonstrate dual specificity *in vivo*. (A) Schematic of the xenograft mouse model in which NSG mice were systemically engrafted with mixed FFluc-eGFP labeled PC3-PSMA (5×10^5 cells) and PC3-MSLN (5×10^5 cells) cells at 1:1 ratio, and treated with two doses of CAR-T cells at day 0 and day 7, respectively (6×10^6 cells each dose, $n = 5$ mice per group). (B, C) Representative tumor BLI images (B) and bioluminescence kinetics (C) at selected time points post T cell injections. (D) Number of human CD45⁺CD3⁺ cells in the peripheral blood collected at day 21 post second T-cell infusion in mice treated as described in (A). Data are shown as individual values and the mean ($n = 5$ mice per group) and are representative of two experiments, $p \geq 0.05$ by one-way ANOVA. (E) Representative antigen expression pattern in the tumor cells isolated from the mice with relapsed tumor in mice treated as described in (A).

Figure 8. Heavy-chain-only-based CAR-T cells express LAG3 and PD-1 upon activation as scFv-based CAR-T cells. (A) Representative flow cytometry plots illustrating Granzyme-B expression in T cells expressing either J591 or PSMA-V_H without co-culture with tumor cells (rest condition). (B-E) Representative flow plots and summary illustrating the kinetics of LAG3 (B,C) and PD-1 (D,E) expression in T cells expressing either J591 or PSMA-V_H cocultured overnight with the tumor cell line expressing PSMA (PC3-PSMA-eGFP) at E:T ratio of 1:2. Data are representative of 4 experiments. $**p < 0.01$ Two-way ANOVA.

Figure 9. Transduction efficiency of MSLN-scFv and MSLN-V_H CARs. (A,B) Representative flow pots (A) and summary (B) illustrating MSLN-scFv and MSLN-V_H expression in T cells. The CD19-specific CAR (CD19) and non-transduced T cells (NT) were used as positive and negative controls, respectively. $****p < 0.0001$, One-way ANOVA.

Figure 10. Expression of MSLN in Aspc-1, PC3 and engineered PC3 cells. Representative flow cytometry histograms showing the expression of MSLN in Aspc-1, PC3 and PC3 cells engineered with retroviral vector to express MSLN.

Figure 11. Bispecific heavy-chain-only-based CAR-T cells demonstrate dual specificity. (A) Representative flow cytometry plots showing coculture of PC3-PSMA-eGFP (PSMA target) and Aspc-1-eGFP (MSLN target) tumor cells with CD19.CAR, PSMA-V_H.CAR, MSLN-V_H.CAR and PSMA-V_H/MSLN-V_H.CAR T cells at E:T ratio of 1:5 for 6 days. At the end of co-culture, cells were collected to enumerate T cells (CD3) and tumor cells (GFP), respectively by flow cytometry. (B) Summary of coculture experiments illustrated in (A); error bars represent SD, ($n = 4$). $****p < 0.0001$, Two-way ANOVA. (C,D) IFN- γ (C) and IL-2 (D) released in the coculture supernatant of the experiments illustrated in (A) as measured by ELISA; error bars represent SD, ($n = 4$). $****p < 0.0001$, Two-way ANOVA.

Figure 12. Phenotypic characterization of T cells in the peripheral blood of mice in the PC3-PSMA and PC3-MSLN Mixed tumor model. CAR-T cells in the peripheral blood at day 21 post second T-cell infusion were identified by the expression of CD45 and CD3 by flow cytometry. PD1 (A), TIM3 (B), CD45RA and CCR7 expression (C) were examined, error bars represent SD, (n = 5). $p \geq 0.05$ by One-way ANOVA.

Figure 13: A)-O) Sequences of V_H single domain antibodies that bind PSMA grouped according to family based on sequence similarity. This figure shows the full length V_H sequence. CDR1, CDR2 and CDR3 are highlighted in bold.

Detailed description

- 10 The present invention will now be further described. In the following passages, different aspects of the invention are defined in more detail. Each aspect so defined may be combined with any other aspect or aspects unless clearly indicated to the contrary. In particular, any feature indicated as being preferred or advantageous may be combined with any other feature or features indicated as being preferred or advantageous.

Generally, nomenclatures used in connection with, and techniques of, cell and tissue culture, pathology, oncology, molecular biology, immunology, microbiology, genetics and protein and nucleic acid chemistry and hybridization described herein are those well-known and commonly used in the art. The methods and techniques of the present disclosure are generally performed according to conventional methods well-known in the art and as described in various general
20 and more specific references that are cited and discussed throughout the present specification unless otherwise indicated. See, *e.g.*, Sambrook *et al.*, Molecular Cloning: A Laboratory Manual (4th ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (2013)). Enzymatic reactions and purification techniques are performed according to manufacturer's specifications, as commonly accomplished in the art or as described herein. The nomenclatures used in connection with, and the laboratory procedures and techniques of, immunology, molecular biology, analytical chemistry, synthetic organic chemistry, and medicinal and pharmaceutical chemistry described herein are those well-known and commonly used in the art. Standard techniques are used for chemical syntheses, chemical analyses, pharmaceutical preparation, formulation, and delivery, and treatment of patients.

- 30 The term V_H domain as used herein refers to an isolated single human V_H domain antibody which is also termed V_H sdAb. A V_H domain is referred to as Humabody® herein. Humabody® is a registered trademark of Crescendo Biologics Ltd. These terms are thus used interchangeably.

The term "isolated" refers to a moiety that is isolated from its natural environment. For example, the term "isolated" refers to a single domain antibody that is substantially free of other single domain antibodies, antibodies or antibody fragments. Moreover, an isolated single domain antibody may be substantially free of other cellular material and/or chemicals.

The terms "single domain antibody", "single variable domain antibody", "single variable heavy chain domain antibody", "single V_H domain antibody", "immunoglobulin single variable domain (ISV)", "immunoglobulin single variable domain antibody", "V_H single domain antibody", "single heavy chain domain", "single variable heavy chain domain", "single V_H domain" or "V_H domain" are all well known in the art and describe the single variable fragment of an antibody that binds to a target antigen.

A single variable "heavy chain domain antibody, single variable heavy chain domain, immunoglobulin single heavy chain variable domain (ISV), human V_H single domain" etc as used herein therefore does not comprise any other parts of a full antibody, but the antigen binding V_H domain only; e.g. it only includes the V_H domain and does not comprise constant heavy chain domains and does not comprise a light chain. A single variable heavy chain domain antibody is capable of specific binding to an antigen in the absence of light chain or other antibody fragments.

V_H domains are small molecules of 12-14 kDa which can be combined into different formats (formatted Humabody®) to give multivalent or multispecific antigen binding domains. V_H domains are robust and are characterised by high affinity and stability in serum³⁰.

Each single V_H domain antibody comprises three CDRs and four FRs, arranged from amino-terminus to carboxy-terminus in the following order: FR1-CDR1-FR2-CDR2-FR3-CDR3-FR4. Thus, in one embodiment of the invention, the domain is a human variable heavy chain (V_H) domain with the following formula FR1-CDR1-FR2-CDR2-FR3-CDR3-FR4.

The term "CDR" refers to the complementarity-determining region within antibody variable sequences. There are three CDRs in each of the variable regions of the heavy chain and the light chain, which are designated CDR1, CDR2 and CDR3, for each of the variable regions. The term "CDR set" refers to a group of three CDRs that occur in a single variable region capable of binding the antigen. The exact boundaries of these CDRs have been defined differently according to different systems. The numbering system described by Kabat is used herein.

The terms "Kabat numbering", "Kabat definitions" and "Kabat labeling" are used interchangeably herein. These terms, which are recognized in the art, refer to a system of

numbering amino acid residues which are more variable (*i.e.*, hypervariable) than other amino acid residues in the heavy and light chain variable regions of an antibody, or an antigen binding portion thereof (Kabat *et al.*, (1971) Ann. NY Acad. Sci. 190:382-391 and Kabat, *et al.*, (1991) Sequences of Proteins of Immunological Interest, Fifth Edition, U.S. Department of Health and Human Services, NIH Publication No. 91-3242).

As used herein, the V_H domain is a human V_H domain. The term "a human V_H domain" includes a V_H domain that is derived from or based on a human V_H domain amino acid or nucleic acid sequence. Thus, the term includes variable heavy chain regions derived from human germline immunoglobulin sequences. The V_H domain can be produced using known methods in the art.

- 10 For example, the term "human V_H domain" includes V_H domains that are isolated from transgenic mice expressing human immunoglobulin V genes, in particular in response to an immunisation with an antigen of interest, for example as described in WO 2016/062990. Such domains are preferably fully human. In one embodiment, a human V_H domain can also include a V_H domain that is derived from or based on a human V_H domain amino acid or nucleic acid sequence encoding such V_H domain. Thus, the term includes variable heavy chain regions derived from or encoded by human germline immunoglobulin sequences. A substantially human V_H domain or V_H domain that is derived from or based on a human V_H domain may include amino acid residues not encoded by human germline immunoglobulin sequences (e.g., mutations introduced *in vitro*, e.g. by random or site-specific mutagenesis, or introduced by
- 20 somatic mutation *in vivo*). The term "human V_H domain" therefore also includes a substantially human V_H domain wherein one or more amino acid residue has been modified. For example, a substantially human V_H domain the V_H domain may include up to 10, for example 1, 2, 3, 4 or 5 amino acid modifications compared to a fully human sequence.

- However, the term "human V_H domain" or "substantially human V_H domain", as used herein, is not intended to include antibodies in which CDR sequences derived from the germline of another mammalian species, such as a mouse, have been grafted onto human framework sequences. Preferably, the term "human V_H domain", as used herein, is also not intended to include camelized V_H domains, that is human V_H domains that have been specifically modified, for example *in vitro* by conventional mutagenesis methods to select predetermined positions
- 30 in the V_H domains sequence and introduce one or more point mutation at the predetermined position to change one or more predetermined residue to a specific residue that can be found in a camelid V_{HH} domain.

As used herein, the term V_H or "variable domain" refers to immunoglobulin variable domains defined by Kabat *et al.*, as referenced above.

A V_H domain is the smallest antigen binding fragment. The term "antibody" broadly refers to any immunoglobulin (Ig) molecule, or antigen binding portion thereof, comprised of four polypeptide chains, two heavy (H) chains and two light (L) chains, or any functional fragment, mutant, variant, or derivation thereof, which retains the essential epitope binding features of an Ig molecule. Such mutant, variant, or derivative antibody formats are known in the art. In a full-length antibody, each heavy chain is comprised of a heavy chain variable region (abbreviated herein as HCVR or V_H) and a heavy chain constant region. The heavy chain constant region is comprised of three domains, C_H1, C_H2 and C_H3. Each light chain is comprised of a light chain variable region (abbreviated herein as LCVR or V_L) and a light chain constant region. The light chain constant region is comprised of one domain, C_L. The V_H and V_L regions can be further subdivided into regions of hypervariability, termed complementarity determining regions (CDR), interspersed with regions that are more conserved, termed framework regions (FR). Each V_H and V_L is composed of three CDRs and four FRs, arranged from amino-terminus to carboxy-terminus in the following order: FR1, CDR1, FR2, CDR2, FR3, CDR3, FR4. Immunoglobulin molecules can be of any type (e.g., IgG, IgE, IgM, IgD, IgA and IgY), class (e.g., IgG 1, IgG2, IgG 3, IgG4, IgA1 and IgA2) or subclass.

An antibody fragment as used herein is a portion of an antibody, for example as F(ab')₂, Fab (Fragment, antibody), scFv (single chain variable chain fragments), single domain antibodies (dAbs), Fv, sFv, and the like. Functional fragments of a full-length antibody retain the target specificity of a full length antibody. Recombinant functional antibody fragments have been used to develop therapeutics as an alternative to therapeutics based on mAbs. scFv fragments (~25kDa) consist of the two variable domains, V_H and V_L. Naturally, V_H and V_L domains are non-covalently associated via hydrophobic interaction and tend to dissociate. However, stable fragments can be engineered by linking the domains with a hydrophilic flexible linker to create a single chain Fv (scFv).

MSLN is a glycoprotein anchored to the plasma membrane by a glycosylphosphatidylinositol (GPI) domain. It is initially synthesized as a 69 kDa cell-surface protein. After cleavage of the amino terminus by the furin protease, a 40-kDa C-terminal fragment remains attached to the membrane and a soluble 32-kDa N-terminal fragment, named MPF (megakaryocyte potentiating factor), is released. A soluble form of MSLN has also been detected in the sera of patients with solid tumors, which is referred to as soluble MSLN-related protein (SMRP).

High mRNA expression of mesothelin is found in mesothelioma, lung, ovarian, breast and pancreatic adenocarcinomas. Mesothelin over-expression has also been noted in some other human cancers, including squamous cell carcinomas of different sites such as cervix, lung and head and neck carcinomas, endometrial adenocarcinomas, colorectal, gastric, and

esophageal cancers (Morello et al Mesothelin-Targeted cars: driving T cells to solid tumors. Cancer Discov 2016).

As used herein, the terms mesothelin or MSLN refer to the 40-kDa protein, mesothelin, which is anchored at the cell membrane by a glycosylphosphatidyl inositol (GPI) linkage and its soluble form that circulates in the serum of cancer patients. MSLN contains N-glycosylation sites. For example, the term refers to a human mesothelin of GenBank accession number AAH03512.1.

The human mesothelin protein sequence and nucleic acid sequences are shown below.

SEQ ID NO. 1 human MSLN amino acid sequence

10 MALPTARPLLGSCGTPALGSLLFLLFSLGWWQPSRTLGETGQEAAPLDGVLANPPNISSLS
 PRQLLGFPCEVSGSLSTERVRELAVALAQKNVKLSTEQRLCLAHRLSEPPEDLDALPLDLLL
 FLNPDAFSGPQACTRFFSRITKANVDLLPRGAPERQRLLPAALACWGVRSLLSEADVRAL
 GGLACDLPGRFVAESADEVLLPRLVSCPGPLDQDQQAARAALQGGGPPYGPSTWSVST
 MDALRGLLPVLGQPIIRSIPQGIVAAWRQRSSRDPSWRQPRTLRLPRFRREVEKTACPSGK
 KAREIDESLIFYKKWELEACVDAALLATQMDRVNAIPFTYEQLDVLKHKLDELYPQGYPESVI
 QHLGYLFLKMSPEDIRKWNVTSLETALKALLEVNKGHEMSPQAPRRPLPQVATLIDRFVKGR
 GQLDKDTLDTLAFYPGYLCSLSPEELSSVPPSSIWAVRPQDLDTCDPRQLDVLYPKARLAF
 QNMNGSEYFVKIQSFLGGAPTEDLKALSQQNVSMDLATFMKLRTDAVLPLTVAEVQKLLGP
 HVEGLKAEERHRPVRDWILRQRQDDLDLGLGLQGGIPNGYLVLDLSMQEALSGTPCLLGP
 20 GPVLTVLALLLASTLA

SEQ ID NO. 2 human MSLN nucleic acid sequence

ATGGCCTTGCCAACGGCTCGACCCCTGTTGGGGTCCTGTGGGACCCCCGCCCTCGGC
 AGCCTCCTGTTCTGCTCTTCAGCCTCGGATGGGTGCAGCCCTCGAGGACCCTGGCTG
 GAGAGACAGGGCAGGAGGCTGCGCCCCCTGGACGGAGTCCTGGCCAACCCACCTAACA
 TTTCCAGCCTCTCCCCTCGCCAACTCCTTGGCTTCCCGTGTGCGGAGGTGTCCGGCCT
 GAGCACGGAGCGTGTCCGGGAGCTGGCTGTGGCCTTGGCACAGAAGAATGTCAAGCT
 CTCAACAGAGCAGCTGCGCTGTCTGGCTCACCGGCTCTCTGAGCCCCCGAGGACCTG
 GACGCCCTCCCATTTGACCTGCTGCTATTCTCAACCCAGATGCGTTCTCGGGGCCCC
 AGGCCTGCACCCGTTTCTTCTCCCGCATCACGAAGGCCAATGTGGACCTGCTCCCGAG
 30 GGGGGCTCCCGAGCGACAGCGGCTGCTGCCTGCGGCTCTGGCCTGCTGGGGTGTGC
 GGGGGTCTCTGCTGAGCGAGGCTGATGTGCGGGCTCTGGGAGGCCTGGCTTGCGACC
 TGCCTGGGCGCTTTGTGGCCGAGTCGGCCGAAGTGCTGCTACCCCGGCTGGTGAGCT
 GCCCGGGACCCCTGGACCAGGACCAGCAGGAGGCAGCCAGGGCGGCTCTGCAGGGC

GGGGGACCCCCCTACGGCCCCCGTCGACATGGTCTGTCTCCACGATGGACGCTCTG
 CGGGGCCTGCTGCCCGTGCTGGGCCAGCCCATCATCCGCAGCATCCCGCAGGGGCATC
 GTGGCCGCGTGGCGGCAACGCTCCTCTCGGGACCCATCCTGGCGGCAGCCTGAACGG
 ACCATCCTCCGGCCGCGGTTCCGGCGGGAAGTGGAGAAGACAGCCTGTCCTTCAGGC
 AAGAAGGCCCGCGAGATAGACGAGAGCCTCATCTTCTACAAGAAGTGGGAGCTGGAAG
 CCTGCGTGGATGCGGCCCTGCTGGCCACCCAGATGGACCGCGTGAACGCCATCCCCT
 TCACCTACGAGCAGCTGGACGTCCTAAAGCATAAACTGGATGAGCTCTACCCACAAGGT
 TACCCCGAGTCTGTGATCCAGCACCTGGGCTACCTCTTCCTCAAGATGAGCCCTGAGG
 ACATTCGCAAGTGAATGTGACGTCCCTGGAGACCCTGAAGGCTTTGCTTGAAGTCAAC
 10 AAAGGGCACGAAATGAGTCCTCAGGCTCCTCGGCGGCCCTCCCACAGGTGGCCACC
 CTGATCGACCGCTTTGTGAAGGGAAGGGGCCAGCTAGACAAAGACACCCTAGACACCC
 TGACCGCCTTCTACCCTGGGTACCTGTGCTCCCTCAGCCCCGAGGAGCTGAGCTCCGT
 GCCCCCAGCAGCATCTGGGCGGTGAGGCCCCAGGACCTGGACACGTGTGACCCAAG
 GCAGCTGGACGTCTCTATCCCAAGGCCCGCCTTGCTTTCCAGAACATGAACGGGTCC
 GAATACTTCGTGAAGATCCAGTCCTTCCTGGGTGGGGCCCCCACGGAGGATTTGAAGG
 CGCTCAGTCAGCAGAATGTGAGCATGGACTTGCCACGTTTCATGAAGCTGCGGACGGA
 TCGGGTGCTGCCGTTGACTGTGGCTGAGGTGCAGAACTTCTGGGACCCCCACGTGGAG
 GGCTGAAGGCGGAGGAGCGGCACCGCCCGGTGCGGGACTGGATCCTACGGCAGCG
 GCAGGACGACCTGGACACGCTGGGGCTGGGGCTACAGGGCGGCATCCCCAACGGCTA
 20 CCTGGTCCTAGACCTCAGCATGCAAGAGGCCCTCTCGGGGACGCCCTGCCTCCTAGGA
 CCTGGACCTGTTCTCACCGTCCTGGCACTGCTCCTAGCCTCCACCCTGGCCTGA

Variants of the sequences above are also included. The term variant is define elsewhere herein but includes biologically active variants with at least 90% sequence identity to the sequences shown above. Different isoforms of MSLN are also included. There are 4 different isoforms of mesothelin which are produced by alternative splicing. SEQ ID NO. 1 is isoform 1 which is the canonical sequence.

Isoforms 2, 3 and 4 differ from isoform 1 as follows:

Isoform 2: The sequence of this isoform differs from the canonical sequence as follows:
 Residues 409-416: Missing.

30 Isoform 3: Also known as: SMRP. The sequence of this isoform differs from the canonical
 sequence as follows: Residues 409-416: Missing. 601-630:
 MQEALSGTPCLLGPVLTVLALLLASTLA (SEQ ID NO. 508) →
 VQGGRGGQARAGGRAGGVEVGALSHPSLCRGPLGDALPPRTWTCSHRPGTAPSLHPGLR
 APLPC (SEQ ID NO. 509)

Isoform 4: The sequence of this isoform differs from the canonical sequence as follows:
Residues 44-44: Missing. Residues 409-416: Missing.

The terms "MSLN binding molecule/protein/polypeptide/agent/moiety", "MSLN antigen binding molecule molecule/protein/polypeptide/agent/moiety", "anti- MSLN single domain antibody", "anti- MSLN single immunoglobulin variable domain", "anti- MSLN heavy chain only antibody" or "anti- MSLN antibody" all refer to a molecule capable of specifically binding to the human MSLN antigen. The binding reaction may be shown by standard methods, for example with reference to a negative control test using an antibody of unrelated specificity. Binding is to human MSLN unless otherwise defined.

- 10 A single domain antibody as described herein, "which binds" or is "capable of binding" an antigen of interest, e.g. human MSLN, is one that binds the antigen with sufficient affinity such that the CAR with the single domain antibody is useful as a therapeutic agent in targeting a cell or tissue expressing the antigen MSLN as described herein. Binding is to the extracellular domain of MSLN.

In one embodiment, the invention relates to an isolated CAR comprising a V_H single domain antibody that specifically binds to human MSLN wherein the V_H single domain antibody comprises a CDR1 comprising SEQ NO. 4 or a sequence with 1, 2 or 3 amino acid modifications, a CDR2 comprising SEQ NO. 5 or a sequence with 1, 2, 3 or 4 amino acid modifications and a CDR3 comprising SEQ NO. 6 or a sequence with 1, 2 or 3 amino acid
20 modifications.

SEQ ID NO. 3 V_H single domain antibody full length sequence (V_H 1.1, also termed MSLN V_H 1.1) (CDRs underlined)

EITLKESGPTLVKPTQTLTLTCTFSGFSLSTSGLGVGWIRQPPGKALEWLALIYWND DKRYR
PSLKNRLTI AKDTSKNQVVL TMTNMDPVDTARYYCAHYSTSSETAFDI RGQGTMTVTVSS

The CDRs of SEQ ID NO. 3 are as follows:

CDR1 SEQ ID NO. 4 : TSGLGVG

CDR2 SEQ ID NO. 5 : LIYWND DKRYRPSLKN

CDR3 SEQ ID NO. 6 : YNTSSETAFDI

- In one embodiment, the V_H single domain antibody comprises or consists of SEQ ID NO. 3 or
30 a variant thereof. In one embodiment, the V_H single domain antibody may be a variant of SEQ

ID NO. 3 having one or more amino acid substitutions, deletions, insertions or other modifications.

In one embodiment, the variant ($V_H1.2$) has a substitution of S to N in CDR3 and the CDR3 sequence is: SEQ ID NO. 7: YNTSSETAFDI

A variant as used herein retains a biological function of the single domain antibody, that is binding to the target antigen (e.g. MSLN) and, thus, the variant antibody or antibody fragment thereof can be sequence engineered. Modifications may include one or more substitution, deletion or insertion of one or more codons encoding the single domain antibody or polypeptide that results in a change in the amino acid sequence as compared with the native sequence provided that the CDRs are as defined above.

Amino acid substitutions in variants as described herein can be the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, such as the replacement of a leucine with a serine, i.e., conservative amino acid replacements. Substitutions, insertions, additions or deletions in the framework region may optionally be in the range of about 1 to 25 or 1 to 50, for example 1 to 5, 1 to 10, 1 to 15, 1 to 20 amino acids, for example 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 amino acids.

In one embodiment, the modification (amino acid substitutions, insertion, addition or deletion) is in HCDR1, HCDR2 and or HCDR3 as defined above.

In one embodiment, variations are only in the framework sequences and the V_H single domain antibody comprises CDR1, 2 and 3 (e.g. SEQ ID Nos 4, 5 and 6 or 4, 5 and 7).

The variation allowed may be determined by systematically making insertions, deletions or substitutions of amino acids in the sequence and testing the resulting variants for activity exhibited by the full-length or mature native sequence.

In one embodiment, the modification is a conservative sequence modification. As used herein, the term "conservative sequence modifications" is intended to refer to amino acid modifications that do not significantly affect or alter the binding characteristics of the antibody containing the amino acid sequence. Such conservative modifications include amino acid substitutions, additions and deletions. Modifications can be introduced into an antibody of the invention by standard techniques known in the art, such as site-directed mutagenesis and PCR-mediated mutagenesis. Conservative amino acid substitutions are ones in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino

acids with basic side chains (e.g., lysine, arginine, histidine), acidic side chains (e.g., aspartic acid, glutamic acid), uncharged polar side chains (e.g., glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine, tryptophan), nonpolar side chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine), beta-branched side chains (e.g., threonine, valine, isoleucine) and aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, histidine). Thus, one or more amino acid residues within the CDR regions of a single domain antibody of the invention can be replaced with other amino acid residues from the same side chain family and the altered antibody can be tested for retained function (i.e. antigen binding) using the functional assays described herein.

- 10 Thus, these amino acid changes can typically be made without altering the biological activity, function, or other desired property of the polypeptide, such as its affinity or its specificity for antigen. In general, single amino acid substitutions in nonessential regions of a polypeptide do not substantially alter biological activity. Furthermore, substitutions of amino acids that are similar in structure or function are less likely to disrupt the polypeptides' biological activity. Abbreviations for the amino acid residues that comprise polypeptides and peptides described herein, and conservative substitutions for these amino acid residues are shown in Table 1 below.

Table 1. Amino Acid Residues and Examples of Conservative Amino Acid Substitutions

Original residue Three letter code, single letter code	Conservative substitution
Alanine, Ala, A	Gly, Ser
Arginine, Arg, R	Lys, His
Asparagine, Asn, N	Gln, His
Aspartic acid Asp, D	Glu, Asn
Cysteine, Cys, C	Ser, Ala
Glutamine, Gln, Q	Asn
Glutamic acid, Glu, E	Asp, Gln
Glycine, Gly, G	Ala
Histidine, His, H	Asn, Gln
Isoleucine, Ile, I	Leu, Val
Leucine, Leu, L	Ile, Val
Lysine, Lys, K	Arg, His
Methionine, Met, M	Leu, Ile, Tyr
Phenylalanine, Phe, F	Tyr, Met, Leu

Proline, Pro, P	Ala
Serine, Ser, S	Thr
Threonine, Thr, T	Ser
Tryptophan, Trp, W	Tyr, Phe
Tyrosine, Tyr, Y	Try, Phe
Valine, Val, V	Ile, Leu

A skilled person will know that there are different ways to identify, obtain and optimise the antigen binding molecules as described herein, including yeast display and *in vitro* and *in vivo* expression libraries. This is further described in the examples. Optimisation techniques known in the art, such as display (e.g., ribosome and/or phage display) and / or mutagenesis (e.g., error-prone mutagenesis) can be used. The invention therefore also comprises sequence optimised variants of the antibodies described herein.

In one embodiment, modifications can be made to decrease the immunogenicity of the single domain antibody. For example, one approach is to revert one or more framework residues to the corresponding human germline sequence. More specifically, a single domain antibody that has undergone somatic mutation may contain framework residues that differ from the germline sequence from which the single domain antibody is derived. Such residues can be identified by comparing the single domain antibody framework sequences to the germline sequences from which the single domain antibody is derived. In one embodiment, all framework sequences are germline sequence.

To return one or more of the amino acid residues in the framework region sequences to their germline configuration, the somatic mutations can be "backmutated" to the germline sequence by, for example, site-directed mutagenesis or PCR-mediated mutagenesis.

Another type of framework modification involves mutating one or more residues within the framework region, or even within one or more CDR regions, to remove T cell epitopes to thereby reduce the potential immunogenicity of the antibody.

In still another embodiment, glycosylation is modified. For example, an aglycosylated antibody can be made (i.e., the antibody lacks glycosylation). Glycosylation can be altered to, for example, increase the affinity of the antibody for antigen. Such carbohydrate modifications can be accomplished by, for example, altering one or more sites of glycosylation within the antibody sequence. For example, one or more amino acid substitutions can be made that result in elimination of one or more variable region framework glycosylation sites to thereby eliminate glycosylation at that site. Such aglycosylation may increase the affinity of the antibody for the antigen.

In one embodiment, the one or more substitution is in the CDR1, 2 or 3 region provided that the CDRs as follows: there may be 1, 2, 3 or more amino acid substitutions in the CDR1, 2 or 3. In another example, there may be 1, 2, 3 amino acid deletions or addition.

In one embodiment, the one or more substitution, addition or deletion is in the framework region. For example, there may be 1 to 20, e.g. 1 to 10 or more amino acid substitutions in the framework regions.

10 Variants can also be defined by reference to sequence identity. Sequence identity as defined herein can be at least 40%, 50%, 60%, 70%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% for example at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence homology. For example, a variant of SEQ ID NO. 3 has at least 75%, 80%, 90% or 95% sequence identity to SEQ ID NO. 3 provided that the CDRs are as defined above.

In one embodiment, one or more non-germline residue in SEQ ID NO. 3 is replaced with a germline residue. Thus, the residue at position 1, 34, 60, 65, 70 and/or 103 is replaced with the germline residue as shown below. In one embodiment, all of these residues are replaced with the germline residue.

Position (Kabat)	residue in SEQ ID NO. 3	residue in germline (VH2-05)
1	E	Q
34	L	V
20 60	R	S
65	N	S
70	A	T
103	R	W

As used herein, the terms sequence "homology" or "identity" generally refers to the percentage of amino acid residues in a sequence that are identical with the residues of the reference polypeptide with which it is compared, after aligning the sequences and in some embodiments after introducing gaps, if necessary, to achieve the maximum percent homology, and not considering any conservative substitutions as part of the sequence identity. Thus, the percent homology between two amino acid sequences is equivalent to the percent identity between
 30 the two sequences. Neither N- or C-terminal extensions, tags or insertions shall be construed

as reducing identity or homology. Methods and computer programs for the alignment are well known. The percent identity between two amino acid sequences can be determined using well known mathematical algorithms.

Sequence identity is commonly defined with reference to the algorithm GAP (Wisconsin GCG package, Accelrys Inc, San Diego USA). GAP uses the Needleman and Wunsch algorithm to align two complete sequences, maximising the number of matches and minimising the number of gaps. Generally, default parameters are used, for example with a gap creation penalty equalling 12 and a gap extension penalty equalling 4. Use of GAP may be preferred but other algorithms may be used, e.g. BLAST, FASTA, the Smith-Waterman algorithm, or the TBLASTN program. In particular, the psi-Blast algorithm may be used. Sequence identity may be defined using the Bioedit, ClustalW algorithm. Alignments can be performed using Snapgene and based on MUSCLE (Multiple Sequence Comparison by Log-Expectation) algorithms.

"Binding affinity" generally refers to the strength of the sum total of non-covalent interactions between a single binding site of a molecule (e.g., an antibody or antigen-binding fragment thereof) and its binding partner (e.g., an antigen). Unless indicated otherwise, as used herein, "binding affinity" refers to intrinsic binding affinity which reflects a 1:1 interaction between members of a binding pair (e.g, antibody or antigen -binding fragment thereof and antigen). The affinity of a molecule X for its partner Y can generally be represented by the dissociation constant (KD).

Affinity can be measured and/or expressed in a number of ways known in the art, including, but not limited to, equilibrium dissociation constant (KD), and equilibrium association constant (KA). The KD is calculated from the quotient of koff/kon, whereas KA is calculated from the quotient of kon/koff. Kon refers to the association rate constant of, e.g, an antibody or antigen-binding fragment thereof to an antigen, and koff refers to the dissociation of, e.g, an antibody or antigen-binding fragment thereof from an antigen. The association rate constant, the dissociation rate constant and the equilibrium dissociation constant are used to represent the binding affinity of an antibody to an antigen. Methods for determining association and dissociation rate constants are well known in the art. The kon and koff can be determined by techniques known to one of ordinary skill in the art, such as BIAcore® or KinExA.

The term "specific binding" or "specifically binds to" or is "specific for" a particular polypeptide or an epitope on a particular polypeptide target as used herein can be exhibited, for example, by a molecule having a KD for the target of at least about 10^{-6} M, alternatively at least about 10^{-7} M, alternatively at least about 10^{-8} M, alternatively at least about 10^{-9} M, alternatively at least about 10^{-10} M, alternatively at least about 10^{-11} M, alternatively at least about 10^{-12} M,

or lower. In one embodiment, the term "specific binding" refers to binding where a molecule binds to a particular polypeptide or epitope on a particular polypeptide without substantially binding to any other polypeptide or polypeptide epitope.

The V_H affinity for MSLN as a single V_H (monomer) recombinant protein is in the nanomolar range, e.g. 20 to 40 nM, e.g. 28-34nM. Measurement may be by Biacore.

The terms "antigen(s)" and "epitope(s)" are well established in the art and refer to the portion of a protein or polypeptide which is specifically recognized by a component of the immune system, e.g. an antibody or a T-cell / B-cell antigen receptor. As used herein, the term "antigen(s)" encompasses antigenic epitopes, e.g. fragments of antigens which are recognized
10 by, and bind to, immune components. Epitopes can be recognized by antibodies in solution, e.g. free from other molecules. Epitopes can also be recognized by T-cell antigen receptors when the epitope is associated with a class I or class II major histocompatibility complex molecule.

The term "epitope" or "antigenic determinant" refers to a site on the surface of an antigen to which an immunoglobulin, antibody or antibody fragment specifically binds. Generally, an antigen has several or many different epitopes and reacts with many different antibodies. The term "specifically" includes linear epitopes and conformational epitopes.

Epitopes within protein antigens can be formed both from contiguous amino acids (usually a linear epitope) or non-contiguous amino acids juxtaposed by tertiary folding of the protein
20 (usually a conformational epitope). Epitopes formed from contiguous amino acids are typically, but not always, retained on exposure to denaturing solvents, whereas epitopes formed by tertiary folding are typically lost on treatment with denaturing solvents. An epitope typically includes at least 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 or 15 amino acids in a unique spatial conformation. Methods for determining what epitopes are bound by a given antibody or antibody fragment (i.e., epitope mapping) are well known in the art and include, for example, immunoblotting and immunoprecipitation assays, wherein overlapping or contiguous peptides from are tested for reactivity with a given antibody or antibody fragment. Competition assays can also be used to determine if a test antibody binds to the same epitope as a reference antibody.

30 The degree of competition can be expressed as a percentage of the reduction in binding. Such competition can be measured using a real time, label-free bio-layer interferometry assay, e.g., on an Octet RED384 biosensor (Pall ForteBio Corp.), ELISA (enzyme-linked immunosorbent assays) or SPR (surface plasmon resonance), HTRF; flow cytometry; fluorescent microvolume assay technology (FMAT) assay, Mirrorball, high content imaging based fluorescent

immunoassays, radioligand binding assays, bio-layer interferometry (BLI), surface plasmon resonance (SPR) and thermal shift assays.

In one aspect, the V_H single domain antibody as described above has one or more of the following properties:

- (a) binds cyno MSLN;
- (b) binds human MSLN with a k_a of about 1.45×10^6 , a k_d of about 4.49×10^{-2} and a K_D of about 3.09×10^{-8} (30.9 nM) using a surface plasmon resonance (SPR) assay on Biacore 8K using immobilised cyno MSLN recombinant protein assay;
- 10 (c) binds human MSLN with a k_a of about $1.15\text{-}1.6 \times 10^6$, a k_d of $3.93\text{-}4.49 \times 10^{-2}$ and a K_D of about $2.8\text{-}3.4 \times 10^{-8}$ (28-34 nM) as measured by BiaCore;
- (d) binds cyno MSLN with a k_a of about 1.07×10^6 , a k_d of about 4.68×10^{-2} and a K_D of about 4.38×10^{-8} (43.8 nM) using a surface plasmon resonance (SPR) assay on Biacore 8K using immobilised cyno MSLN recombinant protein assay;
- (e) has an EC₅₀ of about 1.2 nM as measured using an FMAT assay and/or
- (f) is 96.7% monomer after incubation at 4°C for about 17 hours and 98.1% monomer after incubation at 40°C for about 17 hours.

The V_H single domain antibody as described herein binds human MSLN in the nanomolar range, e.g. 20 to 40 nM, e.g. 28-34 nM as measured by BiaCore. Such binding affinity can be particularly useful in applications such as chimeric antigen receptors it enables specific binding to MSLN expressed on the surface of cells and subsequent activation of the CAR-T cells.

The V_H single domain antibody as described herein also binds cyno MSLN. This is advantageous because of the utility of cynomolgous macaques as the non-human primate species of choice for IND enabling activities.

Multispecific binding molecules

In one aspect, there is provided a binding molecule comprising a single variable heavy chain domain antibody that binds to MSLN as described herein and at least a second moiety that binds to a second antigen, for example a tumor specific antigen. The terms binding agent and binding molecule are used interchangeably herein to refer to such multispecific molecule. The binding molecule may be a fusion protein.

30 The first target and the second target are not the same, i.e. are different targets, e.g., proteins; both may be present on a cell surface. Accordingly, a multispecific, e.g. bispecific binding molecule as described herein can selectively and specifically bind to a cell that expresses (or displays on its cell surface) the first target MSLN and the second target. In one embodiment,

a multispecific polypeptide can bind at least two, at least three, at least four, at least five, at least six, or more targets, wherein the multispecific polypeptide agent has at least two, at least, at least three, at least four, at least five, at least six, or more target binding sites respectively.

In one embodiment, the at least second moiety is a binding molecule that binds to a target of interest, for example selected from an antibody or antibody fragment (e.g., a Fab, F(ab')₂, Fv, a single chain Fv fragment (scFv) or single domain antibody, for example a V_H or V_{HH} domain) or antibody mimetic protein. In one embodiment, the single domain antibody of the invention can be linked to an antibody Fc region or fragment thereof, comprising one or both of C_H2 and C_H3 domains, and optionally a hinge region. In one embodiment, the at least second moiety is a V_H domain.

In one embodiment, the binding molecule is bispecific. Thus, in one aspect, the invention relates to a bispecific molecule comprising a single domain antibody described herein linked to a second functional moiety having a different binding specificity than said single domain antibody.

In one embodiment, the bispecific binding molecule has the following formula: V_H (A)- L-V_H (B) wherein A or B is MSLN. V_H (A) is conjugated to V_H (B), i.e. linked to V_H (B), for example with a peptide linker. L denotes a linker, for example a polypeptide linker.

Each V_H comprises CDR and FR regions. Thus, the binding molecule may have the following formula: FR1(A)-CDR1(A)-FR2(A)-CDR2(A)-FR3(A)-CDR3(A)-FR4(A)-L-FR1(B)-CDR1(B)-FR2(B)-CDR2(BA)-FR3(B)-CDR3(B)-FR4(B). The order of the single V_H domains A and B is not particularly limited, so that, within a polypeptide of the invention, single variable domain A may be located N-terminally and single variable domain B may be located C-terminally, or vice versa wherein A or B is MSLN.

The term "peptide linker" refers to a peptide comprising one or more amino acids. A peptide linker comprises 1 to 44 amino acids, more particularly 2 to 20 amino acids. Peptide linkers are known in the art or are described herein. Suitable, non-immunogenic linker peptides are, for example, linkers that include G and/or S residues, (G₄S)_n, (SG₄)_n or G₄(SG₄)_n peptide linkers, wherein "n" is generally a number between 1 and 10, e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10.

In one embodiment, the peptide is for example selected from GGGGS (SEQ ID NO: 8), GGGGSGGGGS (SEQ ID NO: 9), SGGGSGGGG (SEQ ID NO: 10), GGGGSGGGGSGGGGS (SEQ ID NO: 11), GSGSGSGS (SEQ ID NO: 12), GGSGSGSG (SEQ ID NO: 13), GGSGSG (SEQ ID NO: 14), GGSG (SEQ ID NO: 15) and GGGGSGGGGSGGGGSGGGGSGGGGSGGGGSGGGGS (SEQ ID NO: 16).

The one or more single V_H domain antibody, "binds" or is "capable of binding" an antigen of interest, i.e. targets, antigen with sufficient affinity useful in therapy in targeting a cell or tissue expressing the antigen.

As used herein, the term "target" refers to a biological molecule (e.g., antigen, peptide, polypeptide, protein, lipid, carbohydrate) to which a polypeptide domain which has a binding site can selectively bind. The target can be, for example, an intracellular target (such as an intracellular protein target) or a cell-surface target (such as a membrane protein, e.g., a receptor protein). Preferably, a target is a cell-surface target, such as a cell-surface protein. In one embodiment, the target is a tumor specific antigen.

The target antigen as used herein may be selected from a list including, but not limited to PSMA, Her2, CD123, CD19, CD20, CD22, CD23, CD74, BCMA, CD30, CD33, CD52, EGRF, CECAM6, CAXII, CD24, CEA, cMet, TAG72, MUC1, MUC16, STEAP, EphvIII, FAP, GD2, IL-13Ra2, L1-CAM, PSCA, GPC3, Her3, gpA33, 5T4 and ROR1, CD3, CDE28, CD27, CD40, GITTA, OX40, CD80, CD86, ICOS.

In one embodiment, the binding molecule binds MSLN and PSMA. In one embodiment, the antigen binding domain includes a V_H single domain antibody that binds MSLN as described herein and a V_H single domain antibody that binds specifically PSMA.

Binding to PSMA is to wild type human PSMA (accession NO. Q04609). The sequence for the wild type human PSMA monomer is shown below (SEQ ID NO. 17).

```

1 MWNLLHETDS AVATARRPRW LCAGALVLAG GFFLLGFLFG WFIKSSNEAT NITPKHNMKA
61 FLDELKAENI KKFLYNFTQI PHLAGTEQNF QLAKQIQSQW KEFGLDSVEL AHYDVLLSY
121 NKTHPNYISI INEDGNEIFN TSLFEP PPPG YENVSDIVPP FSAFSPQGMP EGDLYVYNYA
181 RTEDFFKLER DMKINCSGKI VIARYGKVFR GNKVKNLA GAKGVILYSD PADYFAPGVK
241 SYPDGWNLP GGVQRGNILN LNGAGDPLTP GYPANEYAYR RGIAEAVGLP SIPVHPIGY
301 DAQKLEKMG GSAPPDSSWR GSKVPYNVG PGFTGNFSTQ KVKMHIHSTN EVTRIYNVIG
361 TLRGAVEPDR YVILGGHRDS WVFGGIDPQS GAAVVHEIVR SFGTLKKEGW RPRRTILFAS
421 WDAEEFGLLG STEWAEENSRL LQERGVAYI NADSSIEGNY TLRVDCTPLM YSLVHNLTK
481 LKSPDEGFEG KSLYESWTKK SPSPEFSGMP RISKLGSGND FEVFFQRLGI ASGRARYTKN
541 WETNKFSGYP LYHSVYETYE LVEKFYDPMF KYHLTVQVR GGMVFELANS IVPFDICRDY
601 AVVLRKYADK IYSISMHPQ EMKTYSVSFD SLFSVKNFT EIASKFSERL QDFDKSNPIV
661 LRMMNDQLMF LERAFIDPLG LPDRPFYRHV IYAPSSH NKY AGESFPGIYD ALFDIESKVD
721 PSKAWGEVVR QIYVAAFTVQ AAAETLSEVA

```

In one embodiment, the antigen binding domain includes a V_H single domain antibody that bind PSMA which with the following sequence or a variant thereof.

SEQ ID NO. 18 (termed 2.1) full length sequence, CDRs underlined

EVQLVESGGGVVQPGRSLRLSCAASGFSFSGYGMHWVRQAPGKGLEWVAYISYDGSNKY
 YADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAKDPAWGLRLGESSYDFDIWGQ
 GTMVTVSS

The CDRs of SEQ ID NO. 18 are shown below

CDR1 SEQ ID NO. 19: GYGMH

CDR2 SEQ ID NO. 20: YISYDGSNKYYADSVKG

CDR3 SEQ ID NO. 21: DPAWGLRLGESSYDFDI

Suitable V_H single domain antibodies that bind PSMA are described in WO 2017/122017, WO2019/012260 and WO2017/191476, both incorporated herein by reference.

- 10 The variant may have one or more amino acid modification, i.e. a substitution, deletion, addition or addition. In one embodiment, the V_H single domain antibody comprises a CDR1 comprising SEQ NO. 19 or a sequence with 1, 2 or 3 amino acid modification, a CDR2 comprising SEQ NO. 20 or a sequence with 1, 2 or 3 amino acid modification and a CDR3 comprising SEQ NO. 21 or a sequence with 1, 2 or 3 amino acid modification.

In one embodiment, the one or more modification is in the CDR1, 2 or 3 region. For example, there may be 1, 2, 3 or more amino acid modification is in the CDR1, 2 or 3.

In one embodiment, the one or more modification is in the framework region. For example, there may be 1 to 20, e.g. 1 to 10 or more amino acid substitutions framework regions.

- 20 In one embodiment, the V_H single domain antibody that binds PSMA has at least 60%, 70%, 80% or 90% homology thereto, for example 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% homology to SEQ ID NO. 18. In one embodiment, said sequence homology or identity is at least 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99%.

In one embodiment, the V_H single domain antibody that binds PSMA is selected from one of the V_H single domain antibody as shown in Table 2 shown below.

Table 2a Family 1 PSMA binders

Name PSMA	CDR1	CDR2	CDR3	V _H Full length sequence
1.1	SEQ ID NO. 24 SYAMS	SEQ ID NO. 25 SIGENDG TTDYADS VKG	SEQ ID NO. 26 DGVH	SEQ ID NO. 27 EVQLLES GGGLVQP GGSLRLSCAASGF SFSSYAMSWVRQAPGKGLEWVSSIGEN DGT TDYADSVKGRFTISRDN SKSMLYLQ MNSLRVEDTAVYYCVKDG VHWGQGTL VTVSS

1.2	SEQ ID NO. 28 SYAMS	SEQ ID NO. 29 SIGDNNN STEYADS VKG	SEQ ID NO. 30 DGVH	SEQ ID NO. 31 EVQLVESGGGLVQPGGSLRLSCAASGF TFSSYAMSWVRQAPGKGLEWVSSIGDN NNSTEYADSVKGRFTISRDNKSTLYLQ MNSLSAEDTAVYYCVKDG VHWGQGTLV TVSS
1.3	SEQ ID NO. 32 SYAMS	SEQ ID NO. 33 IGDNNNS TDYADSV KG	SEQ ID NO. 34 DGVH	SEQ ID NO. 35 EVQLVESGGGLVQPGGSLRLSCAASGF SFSSYAMSWVRQAPGKGLEWVSSIGDN NNSTDYADSVKGRFTISRDNKSTLYLQ MNSLRAEDTAVYYCVKDG VHWGQGTLV TVSS
1.4	SEQ ID NO. 36 SYAMS	SEQ ID NO. 37 IGDGTYY ADSVKG	SEQ ID NO. 38 DGVH	SEQ ID NO. 39 EVQLVESGGGLVQPGGSLRLSCAASGF TFSSYAMSWVRQAPGKGLEWVSSIGDG TYYADSVKGRFTISRDNKSTLYLQMN SLRAEDTAVYYCAKDGVHWGQGTLVTV SS
1.5	SEQ ID NO. 40 TYAMS	SEQ ID NO. 41 SIGENDR TTYVDS VKG	SEQ ID NO. 42 DGVH	SEQ ID NO. 43 EVQLVESGGGLVQPGGSLRLSCAASGF TFSTYAMSWVRQAPGKGLEWVSSIGEN DRTYYVDSVKGRFTISRDNKSTLYLQ MNSLRAEDTAVYYCAKDGVHWGQGTLV TVSS
1.6	SEQ ID NO. 44 SYAMS	SEQ ID NO. 45 SIGDNNR TTYADSV VKG	SEQ ID NO. 46 DGVH	SEQ ID NO. 47 QVQLVESGGGLVQPGGSLRLSCAASGF TFSSYAMSWVRQAPGKGLEWVSSIGDN NRTTYADSVKGRFTISRDNKSTLYLQ MNSLRAEDTAVYYCAKDGVHWGQGTLV TVSS
1.7	SEQ ID NO. 48 SYAMS	SEQ ID NO. 49 SIGDGTTY YADSVKG	SEQ ID NO. 50 DGVH	SEQ ID NO. 51 EVQLVESGGGLVQPGGSLRLSCAASGF TFSSYAMSWVRQAPGKGLEWVSSIGDG TYYADSVKGRFTISRDNKSTLYLQMN SLRAEDTAVYYCAKDGVHWGQGTLVTV SS
1.8	SEQ ID NO. 52 SYAMS	SEQ ID NO. 53 SIGENDG TTDYADS VKG	SEQ ID NO. 54 DGVH	SEQ ID NO. 55 EVQLLES GGGLVQPGGSLRLSCAASGF SFSSYAMSWVRQAPGKGLEWVSSIGEN DGTTDYADSVKGRFTISRDNKNTLYLQ MNSLRVEDTAVYYCVKDG VHWGQGTLV TVSS
1.9	SEQ ID NO. 56 SYALS	SEQ ID NO. 57 SIGENDG TTDYADS VKG	SEQ ID NO. 58 DGVH	SEQ ID NO. 59 EVQLLES GGGLVQPGGSLRLSCAASGF SFSSYALSWVRQAPGKGLEWVSSIGEN DGTTDYADSVKGRFTISRDNKNTLYLQ MNSLRVEDTAVYYCVKDG VHWGQGTLV TVSS
1.10	SEQ ID NO. 60 SYALS	SEQ ID NO. 61 SIGENNAT TDYADFV KG	SEQ ID NO. 62 DGVH	SEQ ID NO. 63 EVQLLES GGGLVQPGGSLRLSCAASGF SFSSYALSWVRQAPGKGLEWVSSIGEN NATTDYADFVKGRFTISRDNKNTLYLQ MNSLRVEDTAVYYCVKDG VHWGQGTLV TVSS
1.11	SEQ ID NO. 64	SEQ ID NO. 65	SEQ ID NO. 66	SEQ ID NO. 67

	SYALS	SIGENND TTDYADN VKG	DGVH	EVQLLESGGGLVQPGGSLRLSCAASGF SFSSYALSWWRQAPGKGLEWVSSIGEN NDTTDYADNVKGRFTISRDN SKNTLYLQ MNSLRVEDTAVYYCVKDG VHWGQGTL VTVSS
1.12	SEQ ID NO. 68 SYALS	SEQ ID NO. 69 SIGENNAT TDYADAV KG	SEQ ID NO. 70 DGVH	SEQ ID NO. 71 EVQLLESGGGLVQPGGSLRLSCAASGF SFSSYALSWWRQAPGKGLEWVSSIGEN NATTDYADAVKGRFTISRDN SKNTLYLQ MNSLRVEDTAVYYCVKDG VHWGQGTL VTVSS
1.13	SEQ ID NO. 72 SYALS	SEQ ID NO. 73 SIGENNH TTDYAAD VKG	SEQ ID NO. 74 DGVH	SEQ ID NO. 75 EVQLLESGGGLVQPGGSLRLSCAASGF SFSSYALSWWRQAPGKGLEWVSSIGEN NHTTDYAADV KGRFTISRDN SKNTLYLQ MNSLRVEDTAVYYCVKDG VHWGQGTL VTVSS
1.14	SEQ ID NO. 76 SYALS	SEQ ID NO. 77 SIGENNAT TDYADV KG	SEQ ID NO. 78 DGVH	SEQ ID NO. 79 EVQLLESGGGLVQPGGSLRLSCAASGF SFSSYALSWWRQAPGKGLEWVSSIGEN NATTDYADV KGRFTISRDN SKNTLYLQ MNSLRVEDTAVYYCVKDG VHWGQGTL VTVSS
1.15	SEQ ID NO. 80 SYALS	SEQ ID NO. 81 SIGENNH TTDYAAF VKG	SEQ ID NO. 82 DGVH	SEQ ID NO. 83 EVQLLESGGGLVQPGGSLRLSCAASGF SFSSYALSWWRQAPGKGLEWVSSIGEN NHTTDYAAFV KGRFTISRDN SKNTLYLQ MNSLRVEDTAVYYCVKDG VHWGQGTL VTVSS
1.16	SEQ ID NO. 84 SYALS	SEQ ID NO. 85 SIGENNH TTDYADT VKG	SEQ ID NO. 86 DGVH	SEQ ID NO. 87 EVQLLESGGGLVQPGGSLRLSCAASGF SFSSYALSWWRQAPGKGLEWVSSIGEN NHTTDYADTV KGRFTISRDN SKNTLYLQ MNSLRVEDTAVYYCVKDG VHWGQGTL VTVSS
1.17	SEQ ID NO. 88 SYALS	SEQ ID NO. 89 SIGENND TTDYADA VKG	SEQ ID NO. 90 DGVH	SEQ ID NO. 91 EVQLLESGGGLVQPGGSLRLSCAASGF SFSSYALSWWRQAPGKGLEWVSSIGEN NDTTDYADAV KGRFTISRDN SKNTLYLQ MNSLRVEDTAVYYCVKDG VHWGQGTL VTVSS
1.18	SEQ ID NO. 92 SYALS	SEQ ID NO. 93 SIGENNAT TDYAASV KG	SEQ ID NO. 94 DGVH	SEQ ID NO. 95 EVQLLESGGGLVQPGGSLRLSCAASGF SFSSYALSWWRQAPGKGLEWVSSIGEN NATTDYAASV KGRFTISRDN SKNTLYLQ MNSLRVEDTAVYYCVKDG VHWGQGTL VTVSS
1.19	SEQ ID NO. 96 SYALS	SEQ ID NO. 97 SIGENND TTDYAAY VKG	SEQ ID NO. 98 DGVH	SEQ ID NO. 99 EVQLLESGGGLVQPGGSLRLSCAASGF SFSSYALSWWRQAPGKGLEWVSSIGEN NDTTDYAAYV KGRFTISRDN SKNTLYLQ MNSLRVEDTAVYYCVKDG VHWGQGTL VTVSS
1.20	SEQ ID NO. 100 SYALS	SEQ ID NO. 101	SEQ ID NO. 102 DGVH	SEQ ID NO. 103 EVQLLESGGGLVQPGGSLRLSCAASGF SFSSYALSWWRQAPGKGLEWVSSIGEN

		SIGENNH TTDYAAT VKG		NHTTDYAATVKGRFTISRDN SKNTLYLQ MNSLRVEDTAVYYCVKDG VHWGQGTL VTVSS
1.21	SEQ ID NO. 510 SYALS	SEQ ID NO. 511 SIGENDG TTDYADF VKG	SEQ ID NO. 512 DGVH	SEQ ID NO. 513 EVQLLESGGGLVQPGGSLRLSCAASGF SFSSYALSWWRQAPGKGLEWVSSIGEN DGTTDYADFVKGRFTISRDN SKNTLYLQ MNSLRVEDTAVYYCVKDG VHWGQGTL VTVSS
1.22	SEQ ID NO. 514 SYALS	SEQ ID NO. 515 SIGENDG TTDYADN VKG	SEQ ID NO. 516 DGVH	SEQ ID NO. 517 EVQLLESGGGLVQPGGSLRLSCAASGF SFSSYALSWWRQAPGKGLEWVSSIGEN DGTTDYADNVKGRFTISRDN SKNTLYLQ MNSLRVEDTAVYYCVKDG VHWGQGTL VTVSS
1.23	SEQ ID NO. 518 SYALS	SEQ ID NO. 519 SIGENDG TTDYAAD VKG	SEQ ID NO. 520 DGVH	SEQ ID NO. 521 EVQLLESGGGLVQPGGSLRLSCAASGF SFSSYALSWWRQAPGKGLEWVSSIGEN DGTTDYAADVKGRFTISRDN SKNTLYLQ MNSLRVEDTAVYYCVKDG VHWGQGTL VTVSS
1.24	SEQ ID NO. 522 SYALS	SEQ ID NO. 523 SIGENDG TTDYADV VKG	SEQ ID NO. 524 DGVH	SEQ ID NO. 525 EVQLLESGGGLVQPGGSLRLSCAASGF SFSSYALSWWRQAPGKGLEWVSSIGEN DGTTDYADVVKGRFTISRDN SKNTLYLQ MNSLRVEDTAVYYCVKDG VHWGQGTL VTVSS
1.25	SEQ ID NO. 526 SYALS	SEQ ID NO. 527 SIGENDG TTDYAAF VKG	SEQ ID NO. 528 DGVH	SEQ ID NO. 529 EVQLLESGGGLVQPGGSLRLSCAASGF SFSSYALSWWRQAPGKGLEWVSSIGEN DGTTDYAAFVKGRFTISRDN SKNTLYLQ MNSLRVEDTAVYYCVKDG VHWGQGTL VTVSS
1.26	SEQ ID NO. 530 SYALS	SEQ ID NO. 531 SIGENDG TTDYADT VKG	SEQ ID NO. 532 DGVH	SEQ ID NO. 533 EVQLLESGGGLVQPGGSLRLSCAASGF SFSSYALSWWRQAPGKGLEWVSSIGEN DGTTDYADTVKGRFTISRDN SKNTLYLQ MNSLRVEDTAVYYCVKDG VHWGQGTL VTVSS
1.27	SEQ ID NO. 534 SYALS	SEQ ID NO. 535 SIGENDG TTDYADA VKG	SEQ ID NO. 536 DGVH	SEQ ID NO. 537 EVQLLESGGGLVQPGGSLRLSCAASGF SFSSYALSWWRQAPGKGLEWVSSIGEN DGTTDYADAVKGRFTISRDN SKNTLYLQ MNSLRVEDTAVYYCVKDG VHWGQGTL VTVSS
1.28	SEQ ID NO. 538 SYALS	SEQ ID NO. 539 SIGENDG TTDYAAS VKG	SEQ ID NO. 540 DGVH	SEQ ID NO. 541 EVQLLESGGGLVQPGGSLRLSCAASGF SFSSYALSWWRQAPGKGLEWVSSIGEN DGTTDYAASVKGRFTISRDN SKNTLYLQ MNSLRVEDTAVYYCVKDG VHWGQGTL VTVSS
1.29	SEQ ID NO. 542 SYALS	SEQ ID NO. 543 SIGENDG TTDYAAY VKG	SEQ ID NO. 544 DGVH	SEQ ID NO. 545 EVQLLESGGGLVQPGGSLRLSCAASGF SFSSYALSWWRQAPGKGLEWVSSIGEN DGTTDYAAYVKGRFTISRDN SKNTLYLQ

				MNSLRVEDTAVYYCVKDGVHWGQGTL VTVSS
1.30	SEQ ID NO. 546 SYALS	SEQ ID NO. 547 SIGENDG TTDYAAT VKG	SEQ ID NO. 548 DGVH	SEQ ID NO. 549 EVQLLES GGGVLVQP GGSRLRLSCAASGF SFSSYALSWWRQAPGKGLEWVSSIGEN DGTTDYAATVKGRFTISRDN SKNTLYLQ MNSLRVEDTAVYYCVKDGVHWGQGTL VTVSS

Table 2b Family 2 PSMA binders. Note that the PSMA binder identified in SEQ ID 18 belongs to this family.

Name PSMA	CDR1	CDR2	CDR3	V _H Full length sequence
2.3	SEQ ID NO. 104 GYGMH	SEQ ID NO. 105 HISYD GSRN YYAES VKG	SEQ ID NO. 106 DPAWGLR LGELSSYD FDI	SEQ ID NO. 107 EVQLVESGGGVVQPGRSLRLSCAASGFSF SGYGMHWRQAPGKGLEWVAHISYDGSN RYYAESVKGRFTISRDN SKNTLSLQMNSLR AEDTAVYYCAKDPAWGLRLGELSSYDFDI WGQGTMTVTVSS
2.4	SEQ ID NO. 108 GYGMH	SEQ ID NO. 109 VISYD GSRN YYADS VKG	SEQ ID NO. 110 DPAWGLR LGELSSYD FEI	SEQ ID NO. 111 QVTLKESGGGVVQPGRSLKLSCAASGFSF SGYGMHWRQAPGKGLEWVAVISYDGSN RYYADSVKGRFTISRDN SKNTLSLQMNSLR AEDTAVYYCARDPAWGLRLGELSSYDFEI WGQGTMTVTVSS
2.5	SEQ ID NO. 112 GYGMH	SEQ ID NO. 113 VISYD GSRN YYADS VKG	SEQ ID NO. 114 DPAWGLR LGELSSYD FEI	SEQ ID NO. 115 QVQLVESGGGVVQPGRSLRLSCAASGFSF SGYGMHWRQAPGKGLEWVAVISYDGSN RYYADSVKGRFTISRDN SKNTLSLQMNSLR AEDTAVYYCAKDPAWGLRLGELSSYDFEI WGQGTMTVTVSS
2.6	SEQ ID NO. 116 GYGMH	SEQ ID NO. 117 VISYD GSNKY YADSV KG	SEQ ID NO. 118 DPAWGLR LGELSSYK FEI	SEQ ID NO. 119 EVQLVESGGGVVQPGRSLRLSCAASGFSF SGYGMHWRQAPGKGLEWVAVISYDGSN KYYADSVKGRFTISRDN SKNTLYLQMNSLR AEDTAVYYCAKDPAWGLRLGELSSYKFEI WGQGTMTVTVSS
2.7	SEQ ID NO. 120 GYGMH	SEQ ID NO. 121 LISYD GSNKY	SEQ ID NO. 122 DPAWGLR LGEQSSYA FDI	SEQ ID NO. 123 EVQLVESGGGVVQPGRSLRLSCAASGFSF SGYGMHWRQAPGKGLEWVALISYDGSN KYYADSVKGRFTISRDN SKNTLYLQMNSLR

		YADSV KG		AEDTAVYYCAKDPAWGLRLGEQSSYAFDI WGQGTMTVTVSS
2.8	SEQ ID NO. 124 GYGMH	SEQ ID NO. 125 VISYD GSNKY YADSV KG	SEQ ID NO. 126 DPAWGLR LGEQSSYA FE	SEQ ID NO. 127 QVQLVESGGGVVQPGRSLRLSCAASGFSF SGYGMHWRQAPGKGLEWWSVISYDGSN KYYADSVKGRFTISRDN SKNTLYLQMNSLR TEDTAVYYCAKDPAWGLRLGEQSSYAFEI WGQGTMTVTVSS
2.9	SEQ ID NO. 128 GYGMH	SEQ ID NO. 129 VISYD GSNKY YADSV KG	SEQ ID NO. 130 DPAWGLR LGEQSSYA FEI	SEQ ID NO. 131 EVQLLESGGGVVQPGRSLRLSCAASGFSF SGYGMHWRQAPGKGLEWVAVISYDGSN KYYADSVKGRFTISRDN SKNTLYLQMNSLR VEDTAVYYCAKDPAWGLRLGEQSSYAFEI RGQGTTVTVSS
2.10	SEQ ID NO. 132 GYGMH	SEQ ID NO. 133 YISYD GSNR YYADS VKG	SEQ ID NO. 134 DPAWGLR LGESSSYD FDI	SEQ ID NO. 135 EVQLVESGGGVVQPGRSLRLSCAASGFTF SGYGMHWRQAPGKGLEWWAYISYDGSN RYYADSVKGRFTISRDN SKKTL SLQMNSLR AEDTAVYYCAKDPAWGLRLGESSSYDFDI WGQGTMTVTVSS
2.11	SEQ ID NO. 136 GYGLH	SEQ ID NO. 137 YISYD ESNKY YAPSV KG	SEQ ID NO. 138 DPAWGLR LGESSSYD FDI	SEQ ID NO. 139 EVQLVESGGGVVQPGRSLRLSCAASGFSF SGYGLHWRQAPGKGLEWWAYISYDES NK YYAPSVKGRFTISRDN SKNTLYLQMNSLRA EDTAVYYCAKDPAWGLRLGESSSYDFDIW GQGTMTVTVSS
2.12	SEQ ID NO. 140 GYGMH	SEQ ID NO. 141 YISYD KSNKY YADKV KG	SEQ ID NO. 142 DPAWGLR LGESSSYD FDI	SEQ ID NO. 143 EVQLVESGGGVVQPGRSLRLSCAASGFSF SGYGMHWRQAPGKGLEWWAYISYDKSN KYYADKVKGRFTISRDN SKNTLYLQMNSLR AEDTAVYYCAKDPAWGLRLGESSSYDFDI WGQGTMTVTVSS
2.13	SEQ ID NO. 144 GYGLH	SEQ ID NO. 145 YISYD ASNKY YADNV KG	SEQ ID NO. 146 DPAWGLR LGESSSYD FDI	SEQ ID NO. 147 EVQLVESGGGVVQPGRSLRLSCAASGFSF SGYGLHWRQAPGKGLEWWAYISYDASNK YYADNVKGRFTISRDN SKNTLYLQMNSLRA EDTAVYYCAKDPAWGLRLGESSSYDFDIW GQGTMTVTVSS
2.14	SEQ ID NO. 148	SEQ ID NO. 149	SEQ ID NO. 150	SEQ ID NO. 151 EVQLVESGGGVVQPGRSLRLSCAASGFSF SGYGVHWRQAPGKGLEWWAYISYDASN

	GYGVH	YISYD ASNKY YADNV KG	DPAWGLR LGESSSYD FDI	KYYADNVKGRFTISRDN SKNTLYLQMNSLR AEDTAVYYCAKDPAWGLRLGESSSYDFDI WGQGTMTVTVSS
2.15	SEQ ID NO. 152 GYGLH	SEQ ID NO. 153 YISYD KSNKY YADKV KG	SEQ ID NO. 154 DPAWGLR LGESSSYD FDI	SEQ ID NO. 155 EVQLVESGGGVVQPGRSLRLSCAASGFSF SGYGLHWRQAPGKGLEWWAYISYDKSNK YYADKVKGRFTISRDN SKNTLYLQMNSLRA EDTAVYYCAKDPAWGLRLGESSSYDFDIW GQGTMTVTVSS
2.16	SEQ ID NO. 156 GYGAH	SEQ ID NO. 157 YISYD KSNKY YADKV KG	SEQ ID NO. 158 DPAWGLR LGESSSYD FDI	SEQ ID NO. 159 EVQLVESGGGVVQPGRSLRLSCAASGFSF SGYGAHWRQAPGKGLEWWAYISYDKSN KYYADKVKGRFTISRDN SKNTLYLQMNSLR AEDTAVYYCAKDPAWGLRLGESSSYDFDI WGQGTMTVTVSS
2.17	SEQ ID NO. 160 GYGMH	SEQ ID NO. 161 YISYD ASNKY YADNV KG	SEQ ID NO. 162 DPAWGLR LGESSSYD FDI	SEQ ID NO. 163 EVQLVESGGGVVQPGRSLRLSCAASGFSF SGYGMHWRQAPGKGLEWWAYISYDASN KYYADNVKGRFTISRDN SKNTLYLQMNSLR AEDTAVYYCAKDPAWGLRLGESSSYDFDI WGQGTMTVTVSS
2.18	SEQ ID NO. 164 GYGQH	SEQ ID NO. 165 YISYD ASNKY YADNV KG	SEQ ID NO. 166 DPAWGLR LGESSSYD FD	SEQ ID NO. 167 EVQLVESGGGVVQPGRSLRLSCAASGFSF SGYGQHWRQAPGKGLEWWAYISYDASN KYYADNVKGRFTISRDN SKNTLYLQMNSLR AEDTAVYYCAKDPAWGLRLGESSSYDFDI WGQGTMTVTVSS
2.19	SEQ ID NO. 168 GYGFH	SEQ ID NO. 169 YISYD ASNKY YADNV KG	SEQ ID NO. 170 DPAWGLR LGESSSYD FDI	SEQ ID NO. 171 EVQLVESGGGVVQPGRSLRLSCAASGFSF SGYGFHWRQAPGKGLEWWAYISYDASN KYYADNVKGRFTISRDN SKNTLYLQMNSLR AEDTAVYYCAKDPAWGLRLGESSSYDFDI WGQGTMTVTVSS
2.20	SEQ ID NO. 172 GYGMH	SEQ ID NO. 173 IISYDG SNRY ADSVK G	SEQ ID NO. 174 DPAWGLR LGESSSYD FEI	SEQ ID NO. 175 EVQLVESGGGVVQPGRSLRLSCAASGFSF SGYGMHWRQAPGKGLEWWAYISYDGSN RYYADSVKGRFTISRDN SKNTLSLQMNSLR AEDTAVYYCAKDPAWGLRLGESSSYDFEI WGQGTMTVTVSS

2.21	SEQ ID NO. 176 GYGMH	SEQ ID NO. 177 VISYD GNSR YYADS VKG	SEQ ID NO. 178 DPAWGLR LGKLSSYD FEI	SEQ ID NO. 179 QVQLVESGGGVVQPGRSLKLSCAASGFSF SGYGMHWRQAPGKGLEWWAVISYDGSN RYYADSVKGRFTISRDN SKNTLSLQMNSLR AEDTAVYYCAKDPAWGLRLGKLSSYDFEI WGQGTMTVTVSS
2.22	SEQ ID NO. 180 GYGTH	SEQ ID NO. 181 YISYD GSNKY YAAPV KG	SEQ ID NO. 182 DAAWGLR LGESSSYD FDI	SEQ ID NO. 183 EVQLVESGGGVVQPGRSLRLSCAASGFSF SGYGTHWRQAPGKGLEWWAYISYDGSN KYYAAPVKGRFTISRDN SKNTLYLQMNSLR AEDTAVYYCAKDAAWGLRLGESSSYDFDI WGQGTMTVTVSS
2.23	SEQ ID NO. 184 GYGTH	SEQ ID NO. 185 YISYD ESNKY YASSV KG	SEQ ID NO. 186 DRAWGLR LGESSSYD FDI	SEQ ID NO. 187 EVQLVESGGGVVQPGRSLRLSCAASGFSF SGYGTHWRQAPGKGLEWWAYISYDESN KYYASSVKGRFTISRDN SKNTLYLQMNSLR AEDTAVYYCAKDRAWGLRLGESSSYDFDI WGQGTMTVTVSS
2.24	SEQ ID NO. 188 GYGMH	SEQ ID NO. 189 YISYD ESNKY YARLV KG	SEQ ID NO. 190 DTAWGLRL GESSSYDF DI	SEQ ID NO. 191 EVQLVESGGGVVQPGRSLRLSCAASGFSF SGYGMHWRQAPGKGLEWWAYISYDESN KYYARLVKGRFTISRDN SKNTLYLQMNSLR AEDTAVYYCAKDTAWGLRLGESSSYDFDI WGQGTMTVTVSS
2.25	SEQ ID NO. 192 GYGLH	SEQ ID NO. 193 YISYD LSNKY YARGV KG	SEQ ID NO. 194 DVAWGLR LGESSSYD FDI	SEQ ID NO. 195 EVQLVESGGGVVQPGRSLRLSCAASGFSF SGYGLHWRQAPGKGLEWWAYISYDLSNK YYARGVKGRFTISRDN SKNTLYLQMNSLR EDTAVYYCAKDVAWGLRLGESSSYDFDIW GQGTMTVTVSS

Table 2c Family 3 PSMA binders

Name PSMA	CDR1	CDR2	CDR3	V _H Full length sequence
3.1	SEQ ID NO. 196 SYGMH	SEQ ID NO. 197 FMTYD GNSR YYADS VKG	SEQ ID NO. 198 DRIVGGRV PDAFDI	SEQ ID NO. 199 EVQLVESGGGVVQPGRSLRLSCAASGFPLI SYGMHWRQAPGKGLEWWAFMTYDGSN RYYADSVKGRFTISRDN SKNTLYLQMNSLR DEDTALYYCARDIVGGRVPDAFDIWGQG TMVTVSS

3.2	SEQ ID NO. 200 SYGMN	SEQ ID NO. 201 FISYD GSNKY YADSV KKG	SEQ ID NO. 202 DRIVGARV PDAFDI	SEQ ID NO. 203 EVQLVESGGGVVQPGRSLRLSCAASGFPLI SYGMNWRQAPGKGLDWVAFISYDGSNK YYADSVKGRFTISKDNSKNTLYLQMNSLRA EDTAVYYCAKDRIVGARVPDAFDIWGQGT MVTVSS
3.3	SEQ ID NO. 204 SYGMN	SEQ ID NO. 205 FISYD GSNR YYADS VKG	SEQ ID NO. 206 DRIVGARV PDAFDI	SEQ ID NO. 207 EVQLVESGGGVVQPGRSLRLSCAASGFPLI SYGMNWRQAPGKGLEWVAFISYDGSNR YYADSVKGRFTISRDNKNTLYLQMNSLRA EDTAVYYCAKDRIVGARVPDAFDIWGQGT MVTVSS
3.4	SEQ ID NO. 208 SYGMN	SEQ ID NO. 209 FITYD GSNR YYADS VKG	SEQ ID NO. 210 DRIVGARV PDAYDI	SEQ ID NO. 211 EVQLVESGGGAVQPGRSLRLSCAASGFPLI SYGMNWRQAPGKGLDWVAFITYDGSNR YYADSVKGRFTISRDNKNTLYLQMNSLRP EDTAVYYCAKDRIVGARVPDAYDIWGQGT MVTVSS
3.5	SEQ ID NO. 212 SYGMN	SEQ ID NO. 213 FITYD GSNR YYADS VKG	SEQ ID NO. 214 DRIVGARV PDAYDI	SEQ ID NO. 215 QVQLVESGGGVVQPGRSLRLSCAASGFPLI SYGMNWRQAPGKGLDWVAFITYDGSNR YYADSVKGRFTISRDNKNTLYLQMNSLRP EDTAVYYCAKDRIVGARVPDAYDIWGQGT MVTVSS
3.6	SEQ ID NO. 216 SYGMN	SEQ ID NO. 217 FITYD GSNR YYADS VKG	SEQ ID NO. 218 DRIVGARV PDAYDI	SEQ ID NO. 219 EVQLLESGGGVVQPGRSLRLSCAASGFPLI SYGMNWRQAPGKGLDWVAFITYDGSNR YYADSVKGRFTISRDNKNTLYLQMNSLRP EDTAVYYCAKDRIVGARVPDAYDIWGQGT MVTVSS
3.7	SEQ ID NO. 220 SYGMN	SEQ ID NO. 221 FITYD GSNR YYADS VKG	SEQ ID NO. 222 DRIVGARV PDAYDI	SEQ ID NO. 223 QVQLVESGGGLVQPGGSLRLSCAASGFPLI SYGMNWRQAPGKGLDWVAFITYDGSNR YYADSVKGRFTISRDNKNTLHLQMDSLRLP EDTAVYYCAKDRIVGARVPDAYDIWGQGT MVTVSS

3.8	SEQ ID NO. 224 SYGMN	SEQ ID NO. 225 FITYD GSRN YYADS VKG	SEQ ID NO. 226 DRIVGARV PDAYDI	SEQ ID NO. 227 EVQLVESGGGVVQPGRSLRLSCAASGFPLI SYGMNWRQAPGKGLDWWAFITYDGSNR YYADSVKGRFTISRDN SKNTLYLQMNSLRP EDTAVYYCAKDRIVGARV PDAYDIWGQGT LVTVSS
3.9	SEQ ID NO. 228 SYGMN	SEQ ID NO. 229 FISYD GSRN YYADS VKG	SEQ ID NO. 230 DRIVGARV PDAYDI	SEQ ID NO. 231 QVQLVESGGGVVQPGRSLRLSCAASGFPLI ISYGMNWRQAPGKGLEWVAFISYDGSNR YYADSVKGRFTISRDN SKNTLYLQMNSLRP EDTAVYYCAKDRIVGARV PDAYDIWGQGT MVTVSS
3.10	SEQ ID NO. 232 SYGMN	SEQ ID NO. 233 FITYD GSRN YYADS VKG	SEQ ID NO. 234 DRIVGARV PDAYDI	SEQ ID NO. 235 QVQLVESGGGVVQPGRSLRLSCAASGFPLI ISYGMNWRQAPGKGLDWWAFITYDGSNR YYADSVKGRFTISRDN SKNTLH LQMNSLRP EDTAVYYCAKDRIVGARV PDAYDIWGQGT MVTVSS
3.11	SEQ ID NO. 236 SYGMN	SEQ ID NO. 237 FITYD GSRN YYADS VKG	SEQ ID NO. 238 DRIVGARV PDAYDI	SEQ ID NO. 239 EVQLVESGGGVVQPGRSLRLSCAASGFPLI SYGMNWRQAPGKGLDWWAFITYDGSNR YYADSVKGRFTISRDN SKNTLH LQMNSLRP EDTAVYYCAKDRIVGARV PDAYDIWGQGT MVTVSS
3.12	SEQ ID NO. 240 SYGMN	SEQ ID NO. 241 FITYD GSRN YYADS VKG	SEQ ID NO. 242 DRIVGARV PDAYDI	SEQ ID NO. 243 EVQLVESGGGVVQPGRSLRLSCAASGFPLI SYGMNWRQAPGKGLDWWAFITYDGSNR YYADSVKGRFTISRDN SKNTLYLQMNSLRP EDTAVYYCAKDRIVGARV PDAYDIWGQGT MVTVSS
3.13	SEQ ID NO. 244 SYGMN	SEQ ID NO. 245 FITYD GSRN YYADS VKG	SEQ ID NO. 246 DRIVGARV PDAYDI	SEQ ID NO. 247 QVQLVESGGGVVQPGRSLRLSCAASGFPLI ISYGMNWRQAPGKGLDWWAFITYDGSNR YYADSVKGRFTISRDN SKNTLYLQMNSLRP EDTAVYYCAKDRIVGARV PDAYDIWGQGT MVTVSS
3.14	SEQ ID NO. 248	SEQ ID NO. 249	SEQ ID NO. 250 DRIVGARV PDAYDI	SEQ ID NO. 251 EVQLVESGGGVVVRP GGSRLRLSCAASGFPLI SYGMNWRQAPGKGLDWWAFITYDGSNR YYADSVKGRFTISRDN SKNTLH LQMNSLRP

	SYGM N	FITYD GSRN YYADS VKG		EDTAVYYCAKDRIVGARVPDAYDIWGQGT MVTVSS
3.15	SEQ ID NO. 252 SYGM N	SEQ ID NO. 253 FITYD GSRN YYADS VKG	SEQ ID NO. 254 DRIVGARV PDAYDI	SEQ ID NO. 255 EVQLVESGGGLVQPGGSLRLSCAASGFPLI SYGMNWVRQAPGKGLDWVAFITYDGSRN YYADSVKGRFTISRDN SKNTLHLQMNSLRP EDTAVYYCAKDRIVGARVPDAYDIWGQGT MVTVSS
3.16	SEQ ID NO. 256 SYGM N	SEQ ID NO. 257 FITYD GSRN YYADS VKG	SEQ ID NO. 258 DRIVGARV PDAYDI	SEQ ID NO. 259 EVQLLESGGGVVQPGRSLRLSCAASGFPLI SYGMNWVRQAPGKGLDWVAFITYDGSRN YYADSVKGRFTISRDN SKNTLHLQMNSLRP EDTAVYYCAKDRIVGARVPDAYDIWGQGT MVTVSS
3.17	SEQ ID NO. 260 SYGM N	SEQ ID NO. 261 FITYD GSRN YYADS VKG	SEQ ID NO. 262 DRIVGARV PDAYDI	SEQ ID NO. 263 EVQLLESGGGVVQPGRSLRLSCAASGFPLI SYGMNWVRQAPGKGLDWVAFITYDGSRN YYADSVKGRFTISRDN SKNTLYLQMNSLKP EDTAVYYCAKDRIVGARVPDAYDIWGQGT MVTVSS
3.18	SEQ ID NO. 264 SYGM N	SEQ ID NO. 265 FITYD GSRN YYADS VKG	SEQ ID NO. 266 DRIVGARV PDAYDI	SEQ ID NO. 267 EVQLVESGGGVVQPGRSLRLSCAASGFPLI SYGMNWVRQAPGKGLDWVAFITYDGSRN YYADSVKGRFTISRDN SKNTLYLQMNSLKP EDTAVYYCAKDRIVGARVPDAYDIWGQGT MVTVSS
3.19	SEQ ID NO. 268 SYGM H	SEQ ID NO. 269 FMTYD GSRN YYADA VKG	SEQ ID NO. 270 DRIVGGRV PDAFDI	SEQ ID NO. 271 QVQLVESGGGVVQPGRSLRLSCAASGFPL ISYGMHWVRQAPGKGLEWVAFMTYDGSN RYYADAVKGRFTISRDN SKNTLYLQMNSLR AEDTAVYYCARDIVGGRVPDAFDIWGQG TMVTVSS
3.20	SEQ ID NO. 272 SYGM H	SEQ ID NO. 273 FQTYD GSRN YYADA VKG	SEQ ID NO. 274 DRIVGGRV PDAFDI	SEQ ID NO. 275 QVQLVESGGGVVQPGRSLRLSCAASGFPL ISYGMHWVRQAPGKGLEWVAFQTYDGSN RYYADAVKGRFTISRDN SKNTLYLQMNSLR AEDTAVYYCARDIVGGRVPDAFDIWGQG TMVTVSS
3.21	SEQ ID NO. 276 SYGM H	SEQ ID NO. 277 FQTYD GSRN	SEQ ID NO. 278 DRIVGGRV PDAFDI	SEQ ID NO. 279 QVQLVESGGGVVQPGRSLRLSCAASGFPL ISYGMHWVRQAPGKGLEWVAFQTYDGSN RYYADSVKGRFTISRDN SKNTLYLQMNSLR

		YYADS VKG		AEDTAVYYCARDRIVGGRVPDAFDIWGQG TMVTVSS
3.22	SEQ ID NO. 280 SYGM H	SEQ ID NO. 281 FQTYD ASNR YADSV KG	SEQ ID NO. 282 DRIVGGRV PDAFDI	SEQ ID NO. 283 QVQLVESGGGVVQPGRSLRLSCAASGFPL ISYGMHWRQAPGKGLEWVAFQTYDASN RYYADSVKGRFTISRDN SKNTLYLQMNSLR AEDTAVYYCARDRIVGGRVPDAFDIWGQG TMVTVSS
3.23	SEQ ID NO. 284 SYGM H	SEQ ID NO. 285 FQTYD ASNR YADAV KG	SEQ ID NO. 286 DRIVGGRV PDAFDI	SEQ ID NO. 287 QVQLVESGGGVVQPGRSLRLSCAASGFPL ISYGMHWRQAPGKGLEWVAFQTYDASN RYYADAVKGRFTISRDN SKNTLYLQMNSLR AEDTAVYYCARDRIVGGRVPDAFDIWGQG TMVTVSS
3.24	SEQ ID NO. 288 SYGM N	SEQ ID NO. 289 FITYD GSRN YYADS VKG	SEQ ID NO. 290 DRIVGARV PDAYDI	SEQ ID NO. 291 QVQLVESGGGVVQPGRSLRLSCAASGFPL ISYGMNWRQAPGKGLEWVAFITYDGSNR YYADSVKGRFTISRDN SKNTLYLQMNSLR AEDTAVYYCAKDRIVGARVPDAYDIWGQGT MVTVSS

Table 2d Family 4 PSMA binders

Name PSM A	CDR1	CDR2	CDR3	V _H Full length sequence
4.1	SEQ ID NO. 292 SYGMH	SEQ ID NO. 293 VISYDG SNRY ADSVK G	SEQ ID NO. 294 ERIFGVLT PDDFDI	SEQ ID NO. 295 QVQLVESGGGVVQPGRSLRLSCVASGFPP ISYGMHWRQAPGKGREWVAVISYDGSN RYYADSVKGRFTISRDN SKNTLYLQMNSLR PEDTAVYYCAKERIFGVLT PDDFDIWGQGT TVTVSS
4.2	SEQ ID NO. 296 SYGMH	SEQ ID NO. 297 VISYDG SNRY ADSVK G	SEQ ID NO. 298 ERIFGVLT PDDFDI	SEQ ID NO. 299 QVQLVESGGGVVQPGRSLRLSCAASGFPP ISYGMHWRQAPGKGLEWVAVISYDGSNR YYADSVKGRFTISRDN SKNTLYLQMNSLR EDTAVYYCAKERIFGVLT PDDFDIWGQGT TVTVSS
4.3	SEQ ID NO. 300 SYGMH	SEQ ID NO. 301 VISYDG ANRY ADSVK G	SEQ ID NO. 302 ERIFGVLT PDDFEI	SEQ ID NO. 303 EVQLLES GGGVVQPGRSLRLSCAASGFPP ISYGMHWRQAPGKGLEWVAVISYDGSNR YYADSVKGRFTISRDN SKNTLYLQMNSLR EDTAVYYCAKERIFGVLT PDDFEIWGQGT TVTVSS

4.4	SEQ ID NO. 304 SYGMH	SEQ ID NO. 305 VISYDG SNRYYS ADSVK G	SEQ ID NO. 306 ERIFGALTP DDFDI	SEQ ID NO. 307 EVQLVESGGGVVQPGRSLRLSCAASGFTF TSYGMHWVRQAPGKGLEWVAVISYDGSN RYYADSVKGRFTISRDN SKNTLYLQMNSLR PEDTAVYYCAKERIFGALTPDDFDIWGQGT TVTVSS
-----	-------------------------	--	--------------------------------------	--

Table 2e Family 5 PSMA binders

Name PSM A	CDR1	CDR2	CDR3	V _H Full length sequence
5.1	SEQ ID NO. 308 NYGMH	SEQ ID NO. 309 IISYDG NTKYYT DSVKG	SEQ ID NO. 310 GLWPSDV	SEQ ID NO. 311 QVQLVESGGGVVQPGRSLRLSCAASGFT FN NYGMHWVRQAPGKGLEWVAIISYDGN TKYYTDSVKGRFTISRDN SKNTLYLQMNS LRVEDTAVYYCAKGLWPSDVWGQGT TVT VSS
5.2	SEQ ID NO. 312 NYGMH	SEQ ID NO. 313 IISYDG NSKYY TDSVK G	SEQ ID NO. 314 GLWPSDV	SEQ ID NO. 315 EVQLVESGGGVVQPGRSLRLSCAASGFT FN NYGMHWVRQAPGKGLEWVAIISYDGN SKYYTDSVKGRFTISRDN SKNTLYLQMNS LRVEDTAVYYCAKGLWPSDVWGQGT TVT VSS

Table 2f Family 6 PSMA binders

Name PSM A	CDR1	CDR2	CDR3	V _H Full length sequence
6.1	SEQ ID NO. 316 NSGYY WS	SEQ ID NO. 317 FIYYNG SIHYNP SLKS	SEQ ID NO. 318 DGDDYGDY	SEQ ID NO. 319 QVQLQESGPGLVKPSQTLSTCTVSGGSI SNSGYYWSWVRQHPGKDLEWIGFIYYNG SIHYNPSLKS RVIIISVDT SKNQFSLKMNSV TAADTAVYYCARDGDDYGDYLRGQGTLV TVSS
6.2	SEQ ID NO. 320 NSGYY WS	SEQ ID NO. 321 FIYYNG SIHYNP SLKS	SEQ ID NO. 322 DGDDYGDY	SEQ ID NO. 323 QVQLQESGPGLVKPSQTLSTCTVSGGSI SNSGYYWSWVRQHPGKGLEWIGFIYYNG SIHYNPSLKS RVIIISVDT SKNQFSLKMSSVT AADTAVYYCARDGDDYGDYLRGQGTLVT VSS
6.3	SEQ ID NO. 324 NSGYY WS	SEQ ID NO. 325 FIYYNG SIHYNP SLKS	SEQ ID NO. 326 DGDDYGDY	SEQ ID NO. 327 QVQLQESGPGLVKPSQTLSTCTVSGGSI SNSGYYWSWVRQHPGKGLEWIGFIYYNG SIHYNPSLKS RVIIISVDT SKNQFSLKLNSVT AADTAVYYCARDGDDYGDYLRGQGTLVT VSS

6.4	SEQ ID NO. 328 NSGYY WS	SEQ ID NO. 329 FIYYNG SIHYNP SLKS	SEQ ID NO. 330 DGDDYGDY	SEQ ID NO. 331 QVQLQESGPGLVKPSQTLSTCTVSGGSI SNSGYYWSWIRQHPGKGLEWIGFIYYNG SIHYNPSLKSRVVISVDTSKNQFSLKLSSVT AADTAVYYCARDGDDYGDYLRGQGTLVT VSS
6.5	SEQ ID NO. 332 NSGYY WS	SEQ ID NO. 333 FIYYNG SIHYNP SLKS	SEQ ID NO. 334 DGDDYGDY	SEQ ID NO. 335 QVQLQESGPGLVKPSQTLSTCTVSGGSI SNSGYYWSWVRQHPGKGLEWIGFIYYNG SIHYNPSLKSRVTISVDTSKNQFSLKMSSV TAADTAVYYCARDGDDYGDYLRGQGTLV TVSS
6.6	SEQ ID NO. 336 NSGYY WS	SEQ ID NO. 337 FIYYNG SIHYNP SLKS	SEQ ID NO. 338 DGDDYGDY	SEQ ID NO. 339 QVQLQESGPGLVKPSQTLSTCTVSGGSI SNSGYYWSWVRQHPGKGLEWIGFIYYNG SIHYNPSLKSRVTISVDTSKNQFSLKLNSV TAADTAVYYCARDGDDYGDYLRGQGTLV TVSS
6.7	SEQ ID NO. 340 NSGYY WS	SEQ ID NO. 341 FIYYNG SIHYNP SLKS	SEQ ID NO. 342 DGDDYGDY	SEQ ID NO. 343 QVQLQESGPGLVKPSQTLSTCTVSGGSI SNSGYYWSWVRQHPGKGLEWIGFIYYNG SIHYNPSLKSRVTISVDTSKNQFSLKLSSV TAADTAVYYCARDGDDYGDYLRGQGTLV TVSS

Table 2g Family 7 PSMA binders

Name PSMA	CDR1	CDR2	CDR3	V _H Full length sequence
7.1	SEQ ID NO. 344 SYWMY	SEQ ID NO. 345 NINHDGSEKYYV DSVKG	SEQ ID NO. 346 DSLIVGERGY	SEQ ID NO. 347 EVQLVESGGGLVQPG GSLRLSCAASGFTFS SYWMYWVRQAPGKG LEWWANINHDGSEKY YVDSVKGRFTISRDN KNSLYLQMNSLRAED TAVYYCARDSLIVGER GYWGQGTTLVTVSS
7.2	SEQ ID NO. 348 SYWMY	SEQ ID NO. 349 NINHDGSEKYYV DSVKG	SEQ ID NO. 350 DNLIVGERGY	SEQ ID NO. 351 EVQLVESGGGLVQPG GSLRLSCAASGFTFS SYWMYWVRQAPGKG LEWWANINHDGSEKY YVDSVKGRFTISRDN KNSLYLQMNSLRAED TAVYYCARDNLIVGER GYWGQGTTLVTVSS
7.3	SEQ ID NO. 352 SYWMY	SEQ ID NO. 353	SEQ ID NO. 354	SEQ ID NO. 355

		NINHGGSEKYYV DSVKG	DSLIVGERGY	EVQLVESGGGLVQPG GSLRLSCAASGFTFS SYWMYWRQAPGKG LEWWANINHGGSEKY YVDSVKGRFTISRDN KNSLYLQMNSLRAED TAVYYCARD SLIVGER GYWGQGTLVTVSS
7.4	SEQ ID NO. 356 SYWMY	SEQ ID NO. 357 NINHQQSEKYYV DSVKG	SEQ ID NO. 358 DSLIVGERGY	SEQ ID NO. 359 EVQLVESGGGLVQPG GSLRLSCAASGFTFS SYWMYWRQAPGKG LEWWANINHQQSEKY YVDSVKGRFTISRDN KNSLYLQMNSLRAED TAVYYCARD SLIVGER GYWGQGTLVTVSS
7.5	SEQ ID NO. 360 SYWMY	SEQ ID NO. 361 NINHPGSEKYYV DSVKG	SEQ ID NO. 362 DSLIVGERGY	SEQ ID NO. 363 EVQLVESGGGLVQPG GSLRLSCAASGFTFS SYWMYWRQAPGKG LEWWANINHPGSEKY YVDSVKGRFTISRDN KNSLYLQMNSLRAED TAVYYCARD SLIVGER GYWGQGTLVTVSS
7.6	SEQ ID NO. 364 SYWMY	SEQ ID NO. 365 NINHEGSEKYYV DSVKG	SEQ ID NO. 366 DSLIVGERGY	SEQ ID NO. 367 EVQLVESGGGLVQPG GSLRLSCAASGFTFS SYWMYWRQAPGKG LEWWANINHEGSEKY YVDSVKGRFTISRDN KNSLYLQMNSLRAED TAVYYCARD SLIVGER GYWGQGTLVTVSS
7.7	SEQ ID NO. 368 SYWMY	SEQ ID NO. 369 NINHIGSEKYYV DSVKG	SEQ ID NO. 370 DSLIVGERGY	SEQ ID NO. 371 EVQLVESGGGLVQPG GSLRLSCAASGFTFS SYWMYWRQAPGKG LEWWANINHIGSEKYY VDSVKGRFTISRDN KNSLYLQMNSLRAED TAVYYCARD SLIVGER GYWGQGTLVTVSS
7.8	SEQ ID NO. 372 SYWMY	SEQ ID NO. 373 NINHGDSEKYYV DSVKG	SEQ ID NO. 374 DTLIVGERGY	SEQ ID NO. 375

				EVQLVESGGGLVQPG GSLRLSCAASGFTFS SYWMYWRQAPGKG LEWWANINHDGSEKY YVDSVKGRFTISRDN KNSLYLQMNSLRAED TAVYYCARDTLIVGER GYWGQGTTLTVSS
--	--	--	--	---

Table 2h Family 8 PSMA binders

Name PSMA	CDR1	CDR2	CDR3	V _H Full length sequence
8.1	SEQ ID NO. 376 GYWWS	SEQ ID NO. 377 EINHSGSTNY NPSLKS	SEQ ID NO. 378 GPIPATAIPDAF D	SEQ ID NO. 379 QVQLQQWGAGLLKP SETLSLTCAVYGGSF S GYWWSWIRQPPGKGL EWIGEINHSGSTNYP SLKSRVTISVDTSKNQ FSLKLSSVTAADTAVY YCARGPIPATAIPDAF DIWGQGTMTVTVSS

Table 2i Family 9 PSMA binders

Name PSMA	CDR1	CDR2	CDR3	V _H Full length sequence
9.1	SEQ ID NO. 380 GHYWS	SEQ ID NO. 381 DINHSGSTNYP PSLKS	SEQ ID NO. 382 DYGDSRSLFDY	SEQ ID NO. 383 QVQLQQWGAGLLKPSETLSLT CAVYGGSFSGHYWSWIRQPPGK GLEWIGDINHSGSTNYP SLKSRVTISVDTSKNQFSLKLSSVTAADT AVYYCVRDYGDSRSLFDYWGQG TLTVTVSS

Table 2j Family 10 PSMA binders

Name PSMA	CDR1	CDR2	CDR3	V _H Full length sequence
10.1	SEQ ID NO. 384 SYGMY	SEQ ID NO. 385 FMSYDGSNKY YVDSVKG	SEQ ID NO. 386 GDYDFWSGY PDYD	SEQ ID NO. 387 QVQLVESGGGLVQPG GSLRLSCAASGFTFSS YGMHWVRQAPGKGLE WAFMSYDGSNKYYV DSVKGRFTISRDN SKN TLYLQMNSLRAEDTAV YYCAKGDYDFWSGY DYDMDVWGQGTTTVT SS

Table 2k Family 11 PSMA binders

Name PSMA	CDR1	CDR2	CDR3	V _H Full length sequence
11.1	SEQ ID NO. 388 SYGMY	SEQ ID NO. 389 VISYDGSNKY ADSVKG	SEQ ID NO. 390 GGNALYSSG WPDD	SEQ ID NO. 391 EVQLVESGGGLVKPGGS LRLSCAASGFNLISYGM WVRQAPGKGLEWVAVIS YDGSNKYADSVKGRFT ISRDN SKNTLFLQMNSLR VEDTAVYYCAKGGNALY SSGWPDDGFDIRGQGT MVTVSS

Table 2l Family 12 PSMA binders

Name PSMA	CDR1	CDR2	CDR3	V _H Full length sequence
12.1	SEQ ID NO. 392 NFGMH	SEQ ID NO. 393 VISYDGNSKYYAD TVKG	SEQ ID NO. 394 GLWPPMDV	SEQ ID NO. 395 QVQLVESGGGVVQP GRSLRLSCAASGFTF SNFGMHWARQAPGK GLEWVAVISYDGNSK YYADTVKGRFTISRDN

				SKNTLYLEMNSLRAD DTAVYYCAKGLWPPM DVRGQGTTTVTVSS
--	--	--	--	--

Table 2m Family 13 PSMA binders

Name PSMA	CDR1	CDR2	CDR3	V _H Full length sequence
13.1	SEQ ID NO. 396 DYWMT	SEQ ID NO. 397 NIKQDGSEKYY VDSVKG	SEQ ID NO. 398 DRGGAVALY HNGMDM	SEQ ID NO. 399 EVQLVESGGGSVQPG GSLRLSCAASGFTFSD YWMTWVRQVPGKLE WWANIKQDGSEKYYVD SVKGRFTISRDNKNSL YLQMNSLR AEDTAVYY CARDRGGAVALYHNG MDMGGQGTTTVTVSS

Table 2n Family 14 PSMA binders

Name PSMA	CDR1	CDR2	CDR3	V _H Full length sequence
14.1	SEQ ID NO. 400 SYDIN	SEQ ID NO. 401 WMNPNSGNT GYAQKFQG	SEQ ID NO. 402 GNGPGITGTT DY	SEQ ID NO. 403 KCSWWSLGEVKKPGA SVKVSCASGYTFTSY DINWVRQATGQGLEW MGWMNPNSGNTGYAQ KFQGRVTMTRNTSISTA YMESSLRSED TAVYY CARGNGPGITGTTDYW GQGLTVTVSS

Table 2o Family 15 PSMA binders

Name PSMA	CDR1	CDR2	CDR3	V _H Full length sequence
15.1	SEQ ID NO. 404 DYGMS	SEQ ID NO. 405 GINWNGDRTGY ADSVKG	SEQ ID NO. 406 ENVIVPAATY	SEQ ID NO. 407 EVQLVESGGGVV RPGGSLRLSCAA SGFTFDDYGMS WVRQAPGKGLE WWSGINWNGDRT GYADSVKGRFTIS RDNAKNSLYLQM NSLRAEDTALYY CGRENVIVPAATY WGQGTLVTVSS

As described above, a linker such as polypeptide linker (e.g. (G4S)_n) may be used to link the V_H single domain antibody that binds MSLN with the V_H single domain antibody that binds PSMA.

The affinity of bispecific antigen binding domain to huPSMA may be about 100 to about 250pm, for example 116-213pM. The affinity of the bispecific antigen binding domain to MSLN may be about 16 to 49nM.

- 10 In one embodiment, the single domain antibody or binding agent described above comprises a further moiety to prolong the half-life of the binding molecule. The further moiety may comprise a protein, for example an antibody, or part thereof that binds a serum albumin, e.g., human serum albumin (HSA) or mouse serum albumin (MSA). The further moiety may comprise a V_H domain that binds serum albumin, e.g., human serum albumin (HSA) or a variant thereof such as HSA C34S or mouse serum albumin (MSA).

Increased half life can also be conferred by conjugating the molecule to an antibody fragment, for example a V_H domain that increases half life as disclosed in WO2020/099871 or WO2020/229842).

- 20 The term "half-life" as used herein refers to the time taken for the serum concentration of the amino acid sequence, compound or polypeptide to be reduced by 50%, in vivo, for example due to degradation of the sequence or compound and/or clearance or sequestration of the

sequence or compound by natural mechanisms. Half-life may be increased by at least 1.5 times, preferably at least 2 times, such as at least 5 times, for example at least 10 times or more than 20 times, greater than the half-life of the corresponding V_H single domain antibodies of the invention. For example, increased half-life may be more than 1 hours, preferably more than 2 hours, more preferably more than 6 hours, such as more than 12 hours, or even more than 24, 48 or 72 hours, compared to the corresponding V_H single domain antibodies or fusion protein of the invention. The in vivo half-life of an amino acid sequence, compound or polypeptide of the invention can be determined in any manner known per se, such as by pharmacokinetic analysis. Suitable techniques will be clear to the person skilled in the art. Half life can for example be expressed using parameters such as the t_{1/2}-alpha t_{1/2}-beta and the area under the curve (AUC).

Exemplary modifications

In one embodiment, the anti-MSLN single domain antibody or multivalent binding molecule is labelled with a detectable or functional label. A label can be any molecule that produces or can be induced to produce a signal, including but not limited to fluorophores, fluorescers, radiolabels, enzymes, chemiluminescers, a nuclear magnetic resonance active label or photosensitizers. Thus, the binding may be detected and/or measured by detecting fluorescence or luminescence, radioactivity, enzyme activity or light absorbance.

In still other embodiments, the anti-MSLN single V_H domain antibody or multivalent binding molecule is coupled to at least one therapeutic moiety, such as a drug, an enzyme or a toxin. In one embodiment, the therapeutic moiety is a toxin, for example a cytotoxic radionuclide, chemical toxin or protein toxin.

In another aspect, the anti-MSLN single domain antibody or multivalent binding molecule is modified to increase half-life, for example by a chemical modification, especially by PEGylation, or by incorporation in a liposome.

To generate multivalent binding agents and fusion proteins as described above, two or more polypeptides can be connected by a linker, for example a polypeptide linker. Suitable linkers include for example a linker with GS residues such as (Gly4Ser)_n, where n=from 1 to 10, e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10.

Exemplary methods for making the V_H single domain antibody

In one embodiment, a V_H single domain antibody as described herein is generated from human heavy chain only antibody produced in a transgenic rodent that expresses human heavy chain loci.

One aspect also relates to a method for producing a human heavy chain only antibodies capable of binding human MSLN said method comprising

- a) immunising a transgenic rodent, e.g. mouse with an MSLN antigen wherein said rodent expresses a nucleic acid construct comprising unrearranged human heavy chain V genes and is not capable of making functional endogenous light or heavy chains,
- b) isolating human heavy chain only antibodies.

Further steps can include isolating a V_H domain from said heavy chain only antibody, for example by generating a library of sequences comprising V_H domain sequences from said rodent, e.g. mouse and isolating sequences comprising V_H domain sequences from said libraries.

Another aspect also relates to a method for producing a single V_H domain antibody capable of binding human MSLN said method comprising

- a) immunising a transgenic rodent with an MSLN antigen wherein said rodent, e.g. mouse, expresses a nucleic acid construct comprising unrearranged human heavy chain V genes and is not capable of making functional endogenous light or heavy chains,
- b) generating a library of sequences comprising V_H domain sequences from said rodent, e.g. mouse and
- c) isolating sequences comprising V_H domain sequences from said libraries.

Further steps may include identifying a single V_H domain antibody or heavy chain only antibody that binds to human MSLN, for example by using functional assays as shown in the examples.

Methods for preparing or generating the polypeptides, nucleic acids, host cells, products and compositions described herein using in vitro expression libraries can comprise the steps of:

- a) providing a set, collection or library of nucleic acid sequences encoding amino acid sequences; and
- b) screening said set, collection or library for amino acid sequences that can bind to / have affinity for MSLN and
- c) isolating the amino acid sequence(s) that can bind to / have affinity for MSLN.

In the above method, the set, collection or library of amino acid sequences may be displayed on a phage, phagemid, ribosome or suitable micro-organism (such as yeast), such as to

facilitate screening. Suitable methods, techniques and host organisms for displaying and screening (a set, collection or library of) amino acid sequences will be clear to the person skilled in the art (see for example Phage Display of Peptides and Proteins: A Laboratory Manual, Academic Press; 1st edition (October 28, 1996) Brian K. Kay, Jill Winter, John McCafferty).

Libraries, for example phage libraries, are generated by isolating a cell or tissue expressing an antigen-specific, heavy chain-only antibody, cloning the sequence encoding the V_H domain(s) from mRNA derived from the isolated cell or tissue and displaying the encoded protein using a library. The V_H domain(s) can be expressed in bacterial, yeast or other expression systems.

- 10 Another aspect also relates to an isolated V_H single domain antibody or an isolated heavy chain only antibody comprising a V_H domain binding to MSLN comprising an amino acid product of or derived from a human V_H germline sequence. The heavy chain only antibody may be fully human or comprise mouse sequences.

The transgenic rodent, for example a mouse or rat, may have a reduced capacity to express endogenous antibody genes. Thus, in one embodiment, the rodent has a reduced capacity to express endogenous light and/or heavy chain antibody genes. The rodent may therefore comprise modifications to disrupt expression of endogenous light and/or heavy chain antibody genes so that no functional light and/or heavy chains are produced.

- 20 For example, the rodent is a mouse. The mouse may comprise a non-functional endogenous lambda light chain locus. Thus, the mouse does not make a functional endogenous lambda light chain. In one embodiment, the lambda light chain locus is deleted in part or completely or rendered non-functional through insertion, inversion, a recombination event, gene editing or gene silencing. For example, at least the constant region genes C1, C2 and C3 may be deleted or rendered non-functional through insertion or other modification as described above. In one embodiment, the locus is functionally silenced so that the mouse does not make a functional lambda light chain.

- 30 Furthermore, the mouse may comprise a non-functional endogenous kappa light chain locus. Thus, the mouse does not make a functional endogenous kappa light chain. In one embodiment, the kappa light chain locus is deleted in part or completely or rendered non-functional through insertion, inversion, a recombination event, gene editing or gene silencing. In one embodiment, the locus is functionally silenced so that the mouse does not make a functional kappa light chain.

The mouse having functionally silenced endogenous lambda and kappa L-chain loci may, for example, be made as disclosed in WO 2003/000737, which is hereby incorporated by reference in its entirety.

Furthermore, the mouse may comprise a non-functional endogenous heavy chain locus. Thus, the mouse does not make a functional endogenous heavy chain. In one embodiment, the heavy chain locus is deleted in part or completely or rendered non-functional through insertion, inversion, a recombination event, gene editing or gene silencing. In one embodiment, the locus is functionally silenced so that the mouse does not make a functional heavy chain.

For example, as described in WO 2004/076618 (hereby incorporated by reference in its entirety), all 8 endogenous heavy chain constant region immunoglobulin genes (μ , δ , $\gamma 3$, $\gamma 1$, $\gamma 2a$, $\gamma 2b$, ϵ and α) are absent in the mouse, or partially absent to the extent that they are non-functional, or genes δ , $\gamma 3$, $\gamma 1$, $\gamma 2a$, $\gamma 2b$ and ϵ are absent and the flanking genes μ and α are partially absent to the extent that they are rendered non-functional, or genes μ , δ , $\gamma 3$, $\gamma 1$, $\gamma 2a$, $\gamma 2b$ and ϵ are absent and α is partially absent to the extent that it is rendered non-functional, or δ , $\gamma 3$, $\gamma 1$, $\gamma 2a$, $\gamma 2b$, ϵ and α are absent and μ is partially absent to the extent that it is rendered non-functional. By deletion in part is meant that the endogenous locus gene sequence has been deleted or disrupted, for example by an insertion, to the extent that no functional endogenous gene product is encoded by the locus, *i.e.*, that no functional product is expressed from the locus. In another embodiment, the locus is functionally silenced.

For example, the mouse comprises a non-functional endogenous heavy chain locus, a non-functional endogenous lambda light chain locus and a non-functional endogenous kappa light chain locus. The mouse therefore does not produce any functional endogenous light or heavy chains. Thus, the mouse is a triple knockout (TKO) mouse.

The transgenic mouse may comprise a vector, for example a Yeast Artificial Chromosome (YAC) for expressing a heterologous heavy chain locus. YACs are vectors that can be employed for the cloning of very large DNA inserts in yeast. As well as comprising all three cis-acting structural elements essential for behaving like natural yeast chromosomes (an autonomously replicating sequence (ARS), a centromere (CEN) and two telomeres (TEL)), their capacity to accept large DNA inserts enables them to reach the minimum size (150 kb) required for chromosome-like stability and for fidelity of transmission in yeast cells.

For example, the YAC may comprise multiple human V_H , D and J genes in combination with mouse immunoglobulin constant region genes lacking C_H1 domains, mouse enhancer and regulatory regions.

Alternative methods known in the art may be used for deletion or inactivation of endogenous mouse or rat immunoglobulin genes and introduction of human V_H , D and J genes in combination with mouse immunoglobulin constant region genes lacking C_H1 domains, mouse enhancer and regulatory regions.

Transgenic mice can be created according to standard techniques as illustrated in the examples. The two most characterised routes for creating transgenic mice are via pronuclear microinjection of genetic material into freshly fertilised oocytes or via the introduction of stably transfected embryonic stem cells into morula or blastocyst stage embryos. Regardless of how the genetic material is introduced, the manipulated embryos are transferred to pseudo-pregnant female recipients where pregnancy continues and candidate transgenic pups are born.

10 The main differences between these broad methods are that ES clones can be screened extensively before their use to create a transgenic animal. In contrast, pronuclear microinjection relies on the genetic material integrating to the host genome after its introduction and, generally speaking, the successful incorporation of the transgene cannot be confirmed until after pups are born.

There are many methods known in the art to both assist with and determine whether successful integration of transgenes occurs. Transgenic animals can be generated by multiple means including random integration of the construct into the genome, site-specific integration, or homologous recombination. There are various tools and techniques that can be used to both drive and select for transgene integration and subsequent modification including the use of drug resistance markers (positive selection), recombinases, recombination-mediated cassette exchange, negative selection techniques, and nucleases to improve the efficiency of
20 recombination. Most of these methods are commonly used in the modification of ES cells. However, some of the techniques may have utility for enhancing transgenesis mediated via pronuclear injection.

Further refinements can be used to give more efficient generation of the transgenic line within the desired background. As described above, in preferred embodiments, the endogenous mouse immunoglobulin expression is silenced to permit sole use of the introduced transgene for the expression of the heavy-chain only repertoire that can be exploited for drug discovery. Genetically manipulated mice, for example TKO mice that are silenced for all endogenous immunoglobulin loci (mouse heavy chain, mouse kappa chain and mouse lambda chain) can be used as described above. The transfer of any introduced transgene to this TKO background
30 can be achieved via breeding, (either conventional or with the inclusion of an IVF step to give efficient scaling of the process). However, it is also possible to include the TKO background during the transgenesis procedure. For example, for microinjection, the oocytes may be derived from TKO donors. Similarly, ES cells from TKO embryos can be derived for use in transgenesis. Triple knock-out mice into which transgenes have been introduced are referred to herein as TKO/Tg. In one embodiment, the mouse is as described in WO 2016/062990.

The transgenic rodent described above produces human variable heavy chains which can be isolated and used for the generation of human V_H domains, for example as described in WO 2016/062990, WO2016113556 and Teng et al. Diverse human V_H antibody fragments with bio-therapeutic properties from the Crescendo Mouse. N Biotechnol 2020;55:65–76.

In one embodiment, the mouse is as described in WO2016/062990.

Thus the invention also relates to a genetically modified rodent, e.g. a mouse or a rat which expresses a human heavy chain locus and which has been immunized with a MSLN antigen. The rodent may have the features set out above. The invention also relates to a rodent as described above, e.g. a mouse or a rat which expresses a heavy chain only antibody comprising a human V_H domain that binds to human MSLN. Preferably, said rodent is not capable of making functional endogenous kappa and lambda light and/or heavy chains. The human heavy chain locus is located on a transgene which can be as described above.

The invention also relates to an anti-human MSLN single V_H domain antibody or an anti-human MSLN heavy chain only antibody comprising a human V_H domain or obtained or obtainable from a rodent, preferably a mouse, immunised with a human MSLN antigen and which expresses a human heavy chain locus. Preferably, said rodent is not capable of making functional endogenous kappa and lambda light and/or heavy chains. The human heavy chain locus is located on a transgene which can be as described above.

Polynucleotides and cells

In another aspect, the invention relates to an isolated nucleic acid molecule comprising at least one nucleic acid encoding a single domain antibody as defined above. This includes therefore the nucleic acid encoding a V_H domain that targets MSLN as described herein. The nucleic acid sequence encoding SEQ ID NO. 3 is shown below.

GAGATCACCT TGAAGGAGTC TGGTCCTACG CTGGTGAAAC CCACACAGAC CCTCAC
GCTG ACCTGTACCT TCTCTGGCTT CTCACTCAGC ACTAGTGGAC TGGGTGTGGG CT
GGATCCGT CAGCCCCCAG GAAAGGCCCT AGAATGGCTT GCACTCATTT ATTGGAATG
A TGATAAACGC TACAGACCAT CTCTGAAGAA CAGGCTCACC ATCGCCAAGG ACACC
TCCAA AAACCAGGTG GTCCTTACAA TGACCAACAT GGACCCTGTG GACACAGCCA G
ATATTACTG TGCACATTAT AGCACCTCGT CCGAGACTGC TTTTGATATC CGGGGCCA
AG GGACAATGGT CACCGTCTCC TCA

SEQ ID NO. 22

In one embodiment, the variant has a substitution and the C underlined above at position 117 is replaced with an A.

The nucleic acid sequence encoding SEQ ID NO. 18 is shown below.

GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
 CTCTCCTGTGCAGCCTCTGGATTCTCCTTCAGTGGCTATGGCATGCACTGGGTCCGCCA
 GGCTCCAGGCAAGGGGACTGGAGTGGGTGGCATATATATCATATGATGGAAGTAATAAAT
 ACTATGCAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATTCCAAGAACACG
 CTGTATCTGCAAATGAACAGCCTGAGAGCTGAGGACACGGCTGTGTATTACTGTGCGAA
 AGATCCGGCCTGGGGATTACGTTTGGGGGAGTCATCGTCCTATGATTTTGATATCTGGG
 GCCAAGGGACAATGGTCACCGTCTCCTCA

SEQ ID NO. 23

10 The nucleic acid sequence encoding SEQ ID NO. 18 may be as shown below or a variant thereof.

GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
 CTCTCCTGTGCAGCCTCTGGATTCTCCTTCAGTGGCTATGGCATGCACTGGGTCCGCCA
 GGCTCCAGGCAAGGGGACTGGAGTGGGTGGCATATATATCATATGATGGAAGTAATAAAT
 ACTATGCAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATTCCAAGAACACG
 CTGTATCTGCAAATGAACAGCCTGAGAGCTGAGGACACGGCTGTGTATTACTGTGCGAA
 AGATCCGGCCTGGGGATTACGTTTGGGGGAGTCATCGTCCTATGATTTTGATATCTGGG
 GCCAAGGGACAATGGTCACCGTCTCCTCA

SEQ ID NO. 23

20 Nucleic acids encoding other PSMA V_H domains as described herein can also be included. For example, the nucleic acid may encode one of the V_H amino acid sequences shown in Table 2. Exemplary sequences are listed below.

Family 1

SEQ ID NO. 408 (encodes V_H domain 1.1)

GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTGGGGGGTCCCTGAGA
 CTCTCCTGTGCAGCCTCTGGATTCAGTTTTAGCAGCTATGCCATGAGTTGGGTCCGCCA
 GGCTCCAGGGAAGGGGCTGGAGTGGGTCTCAAGTATTGGTGAGAATGATGGTACCACA
 GACTACGCAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATTCCAAGAGTAT
 GCTGTATCTGCAAATGAACAGCCTGAGAGTCGAGGACACGGCCGTCTATTACTGTGTGA
 AAGATGGTGTCCACTGGGGCCAGGGAACCCTGGTCACCGTCTCCTCA

30 SEQ ID NO. 409 (encodes V_H domain 1.2)

GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCTTGGTACAGCCTGGGGGGTCCCTGAGA
 CTCTCCTGTGCAGCCTCTGGATTCACCTTTAGCAGTTATGCCATGAGCTGGGTCCGCCA
 GGCTCCAGGGAAGGGGCTGGAGTGGGTCTCAAGTATTGGTGATAATAATAATAGCACA
 GAGTACGCAGACTCCGTGAAGGGCCGTTACCATCTCCAGAGACAATTCCAAGAGCA
 CGCTGTATCTGCAAATGAACAGCCTGAGCGCCGAGGACACGGCCGTATATTACTGTGT
 GAAAGATGGTGTCCACTGGGGCCAGGGAACCCTGGTCACTGTCTCTTCA

SEQ ID NO. 410 (encodes V_H domain 1.3)

GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCTTGGTACAGCCTGGGGGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCTCCTTTAGCAGTTATGCCATGAGCTGGGTCCGCCA
GGCTCCAGGGAAGGGACTGGAGTGGGTCTCAAGTATTGGTGATAATAATAATAGCACA
GACTACGCAGACTCCGTGAAGGGCCGGTTCACCATCTCCAGAGACAATTCCAAGAGTA
CGCTGTATCTGCAAATGAACAGCCTGAGAGCCGAGGACACGGCCGTATATTACTGTGT
GAAAGATGGTGTCCACTGGGGCCAGGGAACCCTGGTCACTGTCTCCTCA

SEQ ID NO. 411 (encodes V_H domain 1.4)

10 GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCTTGGTACAGCCTGGGGGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTACCTTTAGCAGTTATGCCATGAGCTGGGTCCGCCA
GGCTCCAGGGAAGGGGCTGGAGTGGGTCTCAAGTATTGGTGATGGAACCACATACTAC
GCAGACTCCGTGAAGGGCCGTTTCACCATCTCCAGAGACAATTCCAAGAGCACGCTGT
ATCTGCAAATGAACAGCCTGAGAGCCGAGGACACGGCCGTATATTACTGTGCGAAAGAT
GGTGTCCACTGGGGCCAGGGAACCCTGGTCACTGTCTCCTCA

SEQ ID NO. 412 (encodes V_H domain 1.5)

GAAGTGCAGCTGGTGGAGTCTGGGGGAGGCTTGGTACAGCCTGGGGGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTACCTTTAGCACTTATGCCATGAGCTGGGTCCGCCA
GGCTCCAGGGAAGGGGCTGGAGTGGGTCTCAAGTATTGGTGAAAATGATCGAACCACA
TACTACGTAGACTCCGTGAAGGGCCGGTTCACCATCTCCAGAGACAATTCCAAGAGCAC
GCTGTATCTGCAAATGAACAGCCTGAGAGCCGAGGACACGGCCGTATATTACTGTGCG
20 AAAGATGGTGTCCACTGGGGCCAGGGAACCCTGGTCACTGTCTCTTCA

SEQ ID NO. 413 (encodes V_H domain 1.6)

CAGGTGCAGCTGGTGGAGTCTGGGGGAGGCTTGGTACAGCCTGGGGGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTACCTTTAGCAGTTATGCCATGAGCTGGGTCCGCCA
GGCTCCAGGGAAGGGGCTGGAGTGGGTCTCAAGTATTGGTGATAATAATAGAACCACA
TACTACGCAGACTCCGTGAAGGGCCGGTTCACCATCTCCAGAGACAATTCCAAGAGCA
CGCTGTATCTGCAAATGAACAGCCTGAGAGCCGAGGACACGGCCGTATATTACTGTGC
GAAAGATGGTGTCCACTGGGGCCAGGGAACCCTGGTCACCGTCTCCTCA

SEQ ID NO. 414 (encodes V_H domain 1.7)

30 GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCTTGGTACAGCCTGGGGGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTACCTTTAGCAGTTATGCCATGAGCTGGGTCCGCCA
GGCTCCAGGGAAGGGGCTGGAGTGGGTCTCAAGTATTGGTGATGGAACCACATACTAC
GCAGACTCCGTGAAGGGCCGGTTCACCATCTCCAGAGACAATTCCAAGAGCACGCTGT
ATCTGCAAATGAACAGCCTGAGAGCCGAGGACACGGCCGTATATTACTGTGCGAAAGAT
GGTGTCCACTGGGGCCAGGGAACCCTGGTCACCGTCTCCTCA

SEQ ID NO. 415 (encodes V_H domain 1.8)

GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTGGGGGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCAAGTTTATAGCAGCTATGCCATGAGTTGGGTCCGCCA
GGCTCCAGGGAAGGGGCTGGAGTGGGTCTCAAGTATTGGTGAGAATGATGGTACCACA
GACTACGCAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATTCCAAGAATAC
GCTGTATCTGCAAATGAACAGCCTGAGAGTCGAGGACACGGCCGTCTATTACTGTGTGA
AAGATGGTGTCCACTGGGGCCAGGGAACCCTGGTCACCGTCTCCTCA

SEQ ID NO. 416 (encodes V_H domain 1.9)

10 GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTGGGGGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCAAGTTTATAGCAGCTATGCCCTCAGTTGGGTCCGCCA
GGCTCCAGGGAAGGGGCTGGAGTGGGTCTCAAGTATTGGTGAGAATAACGATACCACA
GACTACGCAGACAACGTGAAGGGCCGATTACCATCTCCAGAGACAATTCCAAGAATAC
GCTGTATCTGCAAATGAACAGCCTGAGAGTCGAGGACACGGCCGTCTATTACTGTGTGA
AAGATGGTGTCCACTGGGGCCAGGGAACCCTGGTCACCGTCTCCTCA

SEQ ID NO. 417 (encodes V_H domain 1.10)

20 GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTGGGGGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCAAGTTTATAGCAGCTATGCCCTCAGTTGGGTCCGCCA
GGCTCCAGGGAAGGGGCTGGAGTGGGTCTCAAGTATTGGTGAGAATAACGCTACCACA
GACTACGCAGACTTCGTGAAGGGCCGATTACCATCTCCAGAGACAATTCCAAGAATAC
GCTGTATCTGCAAATGAACAGCCTGAGAGTCGAGGACACGGCCGTCTATTACTGTGTGA
AAGATGGTGTCCACTGGGGCCAGGGAACCCTGGTCACCGTCTCCTCA

SEQ ID NO. 418 (encodes V_H domain 1.11)

GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTGGGGGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCAAGTTTATAGCAGCTATGCCCTCAGTTGGGTCCGCCA
GGCTCCAGGGAAGGGGCTGGAGTGGGTCTCAAGTATTGGTGAGAATAACGCTACCACA
GACTACGCAGACGCCGTGAAGGGCCGATTACCATCTCCAGAGACAATTCCAAGAATA
CGCTGTATCTGCAAATGAACAGCCTGAGAGTCGAGGACACGGCCGTCTATTACTGTGTG
AAAGATGGTGTCCACTGGGGCCAGGGAACCCTGGTCACCGTCTCCTCA

SEQ ID NO. 419 (encodes V_H domain 1.12)

30 GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTGGGGGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCAAGTTTATAGCAGCTATGCCCTCAGTTGGGTCCGCCA
GGCTCCAGGGAAGGGGCTGGAGTGGGTCTCAAGTATTGGTGAGAATAACGCTACCACA
GACTACGCAGACGCCGTGAAGGGCCGATTACCATCTCCAGAGACAATTCCAAGAATA

CGCTGTATCTGCAAATGAACAGCCTGAGAGTCGAGGACACGGCCGTCTATTACTGTGTG
AAAGATGGTGTCCACTGGGGCCAGGGAACCCTGGTCACCGTCTCCTCA

SEQ ID NO. 420 (encodes V_H domain 1.13)

GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTGGGGGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCAGTTTTAGCAGCTATGCCCTCAGTTGGGTCCGCCA
GGCTCCAGGGAAGGGGCTGGAGTGGGTCTCAAGTATTGGTGAGAATAACCATAACCACA
GACTACGCAGCCGACGTGAAGGGCCGATTCACCATCTCCAGAGACAATTCCAAGAATA
CGCTGTATCTGCAAATGAACAGCCTGAGAGTCGAGGACACGGCCGTCTATTACTGTGTG
AAAGATGGTGTCCACTGGGGCCAGGGAACCCTGGTCACCGTCTCCTCA

10 SEQ ID NO. 421 (encodes V_H domain 1.14)

GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTGGGGGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCAGTTTTAGCAGCTATGCCCTCAGTTGGGTCCGCCA
GGCTCCAGGGAAGGGGCTGGAGTGGGTCTCAAGTATTGGTGAGAATAACGCTACCACA
GACTACGCAGACGTCGTGAAGGGCCGATTCACCATCTCCAGAGACAATTCCAAGAATAC
GCTGTATCTGCAAATGAACAGCCTGAGAGTCGAGGACACGGCCGTCTATTACTGTGTGA
AAGATGGTGTCCACTGGGGCCAGGGAACCCTGGTCACCGTCTCCTCA

SEQ ID NO. 422 (encodes V_H domain 1.15)

GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTGGGGGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCAGTTTTAGCAGCTATGCCCTCAGTTGGGTCCGCCA
20 GGCTCCAGGGAAGGGGCTGGAGTGGGTCTCAAGTATTGGTGAGAATAACCATAACCACA
GACTACGCAGCCTTCGTGAAGGGCCGATTCACCATCTCCAGAGACAATTCCAAGAATAC
GCTGTATCTGCAAATGAACAGCCTGAGAGTCGAGGACACGGCCGTCTATTACTGTGTGA
AAGATGGTGTCCACTGGGGCCAGGGAACCCTGGTCACCGTCTCCTCA

SEQ ID NO. 423 (encodes V_H domain 1.16)

GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTGGGGGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCAGTTTTAGCAGCTATGCCCTCAGTTGGGTCCGCCA
GGCTCCAGGGAAGGGGCTGGAGTGGGTCTCAAGTATTGGTGAGAATAACCATAACCACA
GACTACGCAGACACCGTGAAGGGCCGATTCACCATCTCCAGAGACAATTCCAAGAATAC
GCTGTATCTGCAAATGAACAGCCTGAGAGTCGAGGACACGGCCGTCTATTACTGTGTGA
30 AAGATGGTGTCCACTGGGGCCAGGGAACCCTGGTCACCGTCTCCTCA

SEQ ID NO. 424 (encodes V_H domain 1.17)

GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTGGGGGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCAGTTTTAGCAGCTATGCCCTCAGTTGGGTCCGCCA

GGCTCCAGGGAAGGGGCTGGAGTGGGTCTCAAGTATTGGTGAGAATAACGATACCACA
GACTACGCAGACGCCGTGAAGGGCCGATTACCATCTCCAGAGACAATTCCAAGAATA
CGCTGTATCTGCAAATGAACAGCCTGAGAGTCGAGGACACGGCCGTCTATTACTGTGTG
AAAGATGGTGTCCACTGGGGCCAGGGAACCCTGGTCACCGTCTCCTCA

SEQ ID NO. 425 (encodes V_H domain 1.18)

GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTGGGGGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCAAGTTTGTAGCAGCTATGCCCTCAGTTGGGTCCGCCA
GGCTCCAGGGAAGGGGCTGGAGTGGGTCTCAAGTATTGGTGAGAATAACGCTACCACA
GACTACGCAGCCTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATTCCAAGAATAC
10 GCTGTATCTGCAAATGAACAGCCTGAGAGTCGAGGACACGGCCGTCTATTACTGTGTGA
AAGATGGTGTCCACTGGGGCCAGGGAACCCTGGTCACCGTCTCCTCA

SEQ ID NO. 426 (encodes V_H domain 1.19)

GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTGGGGGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCAAGTTTGTAGCAGCTATGCCCTCAGTTGGGTCCGCCA
GGCTCCAGGGAAGGGGCTGGAGTGGGTCTCAAGTATTGGTGAGAATAACGATACCACA
GACTACGCAGCCTACGTGAAGGGCCGATTACCATCTCCAGAGACAATTCCAAGAATAC
GCTGTATCTGCAAATGAACAGCCTGAGAGTCGAGGACACGGCCGTCTATTACTGTGTGA
AAGATGGTGTCCACTGGGGCCAGGGAACCCTGGTCACCGTCTCCTCA

SEQ ID NO. 427 (encodes V_H domain 1.20)

20 GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCTTGGTACAGCCTGGGGGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCAAGTTTGTAGCAGCTATGCCCTCAGTTGGGTCCGCCA
GGCTCCAGGGAAGGGGCTGGAGTGGGTCTCAAGTATTGGTGAGAATAACCATAACCACA
GACTACGCAGCCACCGTGAAGGGCCGATTACCATCTCCAGAGACAATTCCAAGAATAC
GCTGTATCTGCAAATGAACAGCCTGAGAGTCGAGGACACGGCCGTCTATTACTGTGTGA
AAGATGGTGTCCACTGGGGCCAGGGAACCCTGGTCACCGTCTCCTCA

SEQ ID NO. 428 (encodes V_H domain 2.3)

GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCAAGTTTGTAGCAGCTATGGCATGCACTGGGTCCGCC
AGGCTCCAGGCAAGGGGCTGGAGTGGGTGGCACATATATCATATGATGGAAGTAATAG
30 ATACTATGCAGAATCCGTGAAGGGCCGATTACCATCTCCAGAGAGAATTCCAAGAACA
CGCTGTCTCTGCAAATGAACAGCCTGAGAGCTGAGGACACGGCTGTGTATTACTGTGC
GAAAGATCCGGCCTGGGGATTACGTTTGGGGGAGTTATCGTCCTATGATTTTGACATTT
GGGGCCAAGGGACAATGGTCACTGTCTCTTCA

SEQ ID NO. 429 (encodes V_H domain 2.4)

CAGGTCACCTTGAAGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAAA
CTCTCCTGTGCAGCCTCTGGATTCTCCTTCAGTGGCTATGGCATGCACTGGGTCCGCCA
GGCTCCAGGCAAGGGGCTGGAGTGGGTGGCAGTTATATCATATGATGGAAGTAATAGA
TACTATGCAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATTCCAAGAACAC
GCTGTCTCTGCAAATGAACAGCCTGAGAGCTGAGGACACGGCTGTGTATTACTGTGCG
AGAGATCCGGCCTGGGGATTACGTTTGGGGGAGTTATCGTCCTATGATTTTGAAATCTG
GGCCAAGGGACAATGGTCACCGTCTCCTCA

SEQ ID NO. 430 (encodes V_H domain 2.5)

10 CAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCTCCTTCAGTGGCTATGGCATGCACTGGGTCCGCCA
GGCTCCAGGCAAGGGGCTGGAGTGGGTGGCAGTTATATCATATGATGGAAGTAATAGA
TACTATGCAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATTCCAAGAACAC
ACTGTCTCTGCAAATGAACAGCCTGAGAGCTGAGGACACGGCTGTGTATTACTGTGCGA
AAGATCCGGCCTGGGGATTACGTTTGGGGGAGTTATCGTCCTATGATTTTGAAATTTGG
GGCCAAGGGACAATGGTCACCGTCTCTTCA

SEQ ID NO. 431 (encodes V_H domain 2.6)

GAAGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCTCCTTCAGTGGCTATGGCATGCACTGGGTCCGCCA
GGCTCCAGGCAAGGGGCTGGAGTGGGTGGCAGTTATATCATATGATGGAAGTAATAAA
20 TACTATGCAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATTCCAAGAACAC
GCTATATCTGCAAATGAACAGCCTGAGAGCTGAGGACACGGCTGTGTATTACTGTGCGA
AAGATCCGGCCTGGGGATTACGTTTGGGGGAATATCGTCCTATAAATTTGAAATCTGG
GGCCAAGGGACAATGGTCACCGTCTCTTCA

SEQ ID NO. 432 (encodes V_H domain 2.7)

GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCTCCTTCAGTGGCTATGGCATGCACTGGGTCCGCCA
GGCTCCAGGCAAGGGGCTGGAGTGGGTGGCACTTATATCATATGATGGAAGTAATAAAT
ACTATGCAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATTCCAAGAACACG
CTGTATCTGCAAATGAACAGCCTGAGAGCTGAGGACACGGCTGTGTATTACTGTGCGAA
30 AGATCCGGCCTGGGGATTACGTTTGGGGGAGCAATCGTCCTATGCTTTTGATATCTGGG
GCCAAGGGACAATGGTCACCGTCTCCTCA

SEQ ID NO. 433 (encodes V_H domain 2.8)

CAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCTCCTTCAGTGGCTATGGCATGCACTGGGTCCGCCA

GGCTCCAGGCAAGGGGCTGGAGTGGGTGTCAGTTATATCATATGATGGAAGTAATAAAT
ACTATGCAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATCCAAGAACACG
CTGTATCTGCAAATGAACAGCCTGAGAACTGAGGACACGGCTGTGTATTACTGTGCGAA
AGATCCGGCCTGGGGATTACGTTTGGGGGAGCAATCGTCCTATGCTTTTGAAATCTGGG
GCCAAGGTACAATGGTCACCGTCTCCTCA

SEQ ID NO. 434 (encodes V_H domain 2.9)

GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCTCCTTCAGTGGCTATGGCATGCACTGGGTCCGCCA
GGCTCCAGGCAAGGGGCTGGAGTGGGTGGCAGTTATATCATATGATGGAAGTAATAAA
10 TACTATGCAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATCCAAGAACAC
GCTGTATCTGCAAATGAACAGCCTGAGAGTTGAGGACACGGCTGTGTATTACTGTGCGA
AAGATCCGGCCTGGGGATTACGTTTGGGGGAGCAATCGTCCTATGCTTTTGAAATCCGG
GGCCAGGGGACAACGGTCACCGTCTCTTCA

SEQ ID NO. 435 (encodes V_H domain 2.10)

GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTACCTTTAGTGGCTATGGCATGCACTGGGTCCGCCA
GGCTCCAGGCAAGGGGCTGGAGTGGGTGGCATATATATCATATGATGGAAGTAATAGA
TACTATGCAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATCCAAGAAGAC
GCTGTCTCTGCAAATGAACAGCCTGAGAGCTGAGGACACGGCTGTGTATTACTGTGCG
20 AAAGATCCGGCCTGGGGATTACGTTTGGGGGAGTCATCGTCATATGATTTTGATATCTG
GGCCAAGGGACAATGGTCACCGTCTCCTCA

SEQ ID NO. 436 (encodes V_H domain 2.11)

GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCTCCTTCAGTGGCTATGGCCTCCACTGGGTCCGCCA
GGCTCCAGGCAAGGGACTGGAGTGGGTGGCATATATATCATATGACGAGAGTAATAAAT
ACTATGCACCCAGCGTGAAGGGCCGATTACCATCTCCAGAGACAATCCAAGAACACG
CTGTATCTGCAAATGAACAGCCTGAGAGCTGAGGACACGGCTGTGTATTACTGTGCGAA
AGATCCGGCCTGGGGATTACGTTTGGGGGAGTCATCGTCCTATGATTTTGATATCTGGG
GCCAAGGGACAATGGTCACCGTCTCCTCA

30 SEQ ID NO. 437 (encodes V_H domain 2.12)

GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCTCCTTCAGTGGCTATGGCATGCACTGGGTCCGCCA
GGCTCCAGGCAAGGGACTGGAGTGGGTGGCATATATATCATATGATAAGAGTAATAAAT
ACTATGCAGACAAGGTGAAGGGCCGATTACCATCTCCAGAGACAATCCAAGAACACG

CTGTATCTGCAAATGAACAGCCTGAGAGCTGAGGACACGGCTGTGTATTACTGTGCGAA
AGATCCGGCCTGGGGATTACGTTTGGGGGAGTCATCGTCCTATGATTTTGATATCTGGG
GCCAAGGGACAATGGTCACTGTCTCTTCA

SEQ ID NO. 438 (encodes V_H domain 2.13)

GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCTCCTTCAGTGGCTATGGCCTCCACTGGGTCCGCCA
GGCTCCAGGCAAGGGACTGGAGTGGGTGGCATATATATCATATGATGCGAGTAATAAAT
ACTATGCAGACAACGTGAAGGGCCGATTCACCATCTCCAGAGACAATTCCAAGAACACG
CTGTATCTGCAAATGAACAGCCTGAGAGCTGAGGACACGGCTGTGTATTACTGTGCGAA
10 AGATCCGGCCTGGGGATTACGTTTGGGGGAGTCATCGTCCTATGATTTTGATATCTGGG
GCCAAGGGACAATGGTCACTGTCTCTTCA

SEQ ID NO. 439 (encodes V_H domain 2.14)

GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCTCCTTCAGTGGCTATGGCGTGCACTGGGTCCGCC
AGGCTCCAGGCAAGGGACTGGAGTGGGTGGCATATATATCATATGATGCGAGTAATAAA
TACTATGCAGACAACGTGAAGGGCCGATTCACCATCTCCAGAGACAATTCCAAGAACAC
GCTGTATCTGCAAATGAACAGCCTGAGAGCTGAGGACACGGCTGTGTATTACTGTGCGA
AAGATCCGGCCTGGGGATTACGTTTGGGGGAGTCATCGTCCTATGATTTTGATATCTGG
GGCCAAGGGACAATGGTCACTGTCTCTTCA

20 SEQ ID NO. 440 (encodes V_H domain 2.15)

GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCTCCTTCAGTGGCTATGGCCTCCACTGGGTCCGCCA
GGCTCCAGGCAAGGGACTGGAGTGGGTGGCATATATATCATATGATAAGAGTAATAAAT
ACTATGCAGACAAGGTGAAGGGCCGATTCACCATCTCCAGAGACAATTCCAAGAACACG
CTGTATCTGCAAATGAACAGCCTGAGAGCTGAGGACACGGCTGTGTATTACTGTGCGAA
AGATCCGGCCTGGGGATTACGTTTGGGGGAGTCATCGTCCTATGATTTTGATATCTGGG
GCCAAGGGACAATGGTCACTGTCTCTTCA

SEQ ID NO. 441 (encodes V_H domain 2.16)

GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
30 CTCTCCTGTGCAGCCTCTGGATTCTCCTTCAGTGGCTATGGCGCGCACTGGGTCCGCC
AGGCTCCAGGCAAGGGACTGGAGTGGGTGGCATATATATCATATGATAAGAGTAATAAA
TACTATGCAGACAAGGTGAAGGGCCGATTCACCATCTCCAGAGACAATTCCAAGAACAC
GCTGTATCTGCAAATGAACAGCCTGAGAGCTGAGGACACGGCTGTGTATTACTGTGCGA

AAGATCCGGCCTGGGGATTACGTTTGGGGGAGTCATCGTCCTATGATTTTGATATCTGG
GGCCAAGGGACAATGGTCACTGTCTCTTCA

SEQ ID NO. 442 (encodes V_H domain 2.17)

GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCTCCTTCAGTGGCTATGGCATGCACTGGGTCCGCCA
GGCTCCAGGCAAGGGACTGGAGTGGGTGGCATATATATCATATGATGCGAGTAATAAAT
ACTATGCAGACAACGTGAAGGGCCGATTCACCATCTCCAGAGACAATTCCAAGAACACG
CTGTATCTGCAAATGAACAGCCTGAGAGCTGAGGACACGGCTGTGTATTACTGTGCGAA
AGATCCGGCCTGGGGATTACGTTTGGGGGAGTCATCGTCCTATGATTTTGATATCTGGG
10 GGCCAAGGGACAATGGTCACTGTCTCTTCA

SEQ ID NO. 443 (encodes V_H domain 2.18)

GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCTCCTTCAGTGGCTATGGCCAGCACTGGGTCCGCC
AGGCTCCAGGCAAGGGACTGGAGTGGGTGGCATATATATCATATGATGCGAGTAATAAA
TACTATGCAGACAACGTGAAGGGCCGATTCACCATCTCCAGAGACAATTCCAAGAACAC
GCTGTATCTGCAAATGAACAGCCTGAGAGCTGAGGACACGGCTGTGTATTACTGTGCGA
AAGATCCGGCCTGGGGATTACGTTTGGGGGAGTCATCGTCCTATGATTTTGATATCTGG
GGCCAAGGGACAATGGTCACTGTCTCTTCA

SEQ ID NO. 444 (encodes V_H domain 2.19)

20 GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCTCCTTCAGTGGCTATGGCTTCCACTGGGTCCGCCA
GGCTCCAGGCAAGGGACTGGAGTGGGTGGCATATATATCATATGATGCGAGTAATAAAT
ACTATGCAGACAACGTGAAGGGCCGATTCACCATCTCCAGAGACAATTCCAAGAACACG
CTGTATCTGCAAATGAACAGCCTGAGAGCTGAGGACACGGCTGTGTATTACTGTGCGAA
AGATCCGGCCTGGGGATTACGTTTGGGGGAGTCATCGTCCTATGATTTTGATATCTGGG
GGCCAAGGGACAATGGTCACTGTCTCTTCA

SEQ ID NO. 445 (encodes V_H domain 2.20)

30 GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCAGCTTCAGTGGCTATGGCATGCACTGGGTCCGCC
AGGCTCCAGGCAAGGGGCTGGAGTGGGTGGCAATTATATCATATGATGGAAGTAATAG
ATACTATGCAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAATTCCAAGAACA
CGCTGTCTCTGCAAATGAACAGCCTGAGAGCTGAGGACACGGCTGTGTATTACTGTGC
GAAAGATCCGGCCTGGGGATTACGTTTGGGGGAGTCATCGTCCTATGATTTTGAAATTT
GGGGCCAAGGGACAATGGTCACCGTCTCCTCA

SEQ ID NO. 446 (encodes V_H domain 2.21)

CAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAAA
CTCTCCTGTGCAGCCTCTGGATTCTCCTTCAGTGGCTATGGCATGCACTGGGTCCGCCA
GGCTCCAGGCAAGGGGCTGGAGTGGGTGGCAGTTATATCATATGATGGAAGTAATAGA
TACTATGCAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAATTCCAAGAACAC
GCTGTCTCTACAAATGAACAGCCTGAGAGCTGAGGACACGGCTGTGTATTACTGTGCGA
AAGATCCGGCCTGGGGATTACGTTTGGGGAAATTATCGTCCTATGATTTTGAAATCTGG
GGCCAAGGGACAATGGTCACTGTCTCTTCA

SEQ ID NO. 447 (encodes V_H domain 2.22)

10 GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCTCCTTCAGTGGCTATGGCACGCACTGGGTCCGCC
AGGCTCCAGGCAAGGGACTGGAGTGGGTGGCATATATATCATATGACGGGAGTAATAA
ATACTATGCAGCCCCGGTGAAGGGCCGATTCACCATCTCCAGAGACAATTCCAAGAACA
CGCTGTATCTGCAAATGAACAGCCTGAGAGCTGAGGACACGGCTGTGTATTACTGTGC
GAAAGACGCGGCCTGGGGATTACGTTTGGGGGAGTCATCGTCCTATGATTTTGATATCT
GGGGCCAAGGGACAATGGTCACTGTCTCTTCA

SEQ ID NO. 448 (encodes V_H domain 2.23)

GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCTCCTTCAGTGGCTATGGCACGCACTGGGTCCGCC
20 AGGCTCCAGGCAAGGGACTGGAGTGGGTGGCATATATATCATATGACGAGAGTAATAAA
TACTATGCATCCAGCGTGAAGGGCCGATTCACCATCTCCAGAGACAATTCCAAGAACAC
GCTGTATCTGCAAATGAACAGCCTGAGAGCTGAGGACACGGCTGTGTATTACTGTGCGA
AAGACCGGGCCTGGGGATTACGTTTGGGGGAGTCATCGTCCTATGATTTTGATATCTGG
GGCCAAGGGACAATGGTCACTGTCTCTTCA

SEQ ID NO. 449 (encodes V_H domain 2.24)

GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCTCCTTCAGTGGCTATGGCATGCACTGGGTCCGCCA
GGCTCCAGGCAAGGGACTGGAGTGGGTGGCATATATATCATATGACGAGAGTAATAAAT
ACTATGCAAGGCTGGTGAAGGGCCGATTCACCATCTCCAGAGACAATTCCAAGAACACG
30 CTGTATCTGCAAATGAACAGCCTGAGAGCTGAGGACACGGCTGTGTATTACTGTGCGAA
AGACACGGCCTGGGGATTACGTTTGGGGGAGTCATCGTCCTATGATTTTGATATCTGGG
GCCAAGGGACAATGGTCACTGTCTCTTCA

SEQ ID NO. 450 (encodes V_H domain 2.25)

GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
 CTCTCCTGTGCAGCCTCTGGATTCTCCTTCAGTGGCTATGGCCTCCACTGGGTCCGCCA
 GGCTCCAGGCAAGGGGACTGGAGTGGGTGGCATATATATCATATGACCTGAGTAATAAAT
 ACTATGCAAGGGGGGTGAAGGGCCGATTACCATCTCCAGAGACAATTCCAAGAACAC
 GCTGTATCTGCAAATGAACAGCCTGAGAGCTGAGGACACGGCTGTGTATTACTGTGCGA
 AAGACGTGGCCTGGGGATTACGTTTGGGGGAGTCATCGTCCTATGATTTTGATATCTGG
 GGCCAAGGGACAATGGTCACTGTCTCCTCA

family 3

SEQ ID NO. 451 (encodes V_H domain 3.1)

10 GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
 CTCTCCTGTGCAGCCTCTGGATTCCCCTTAATTAGCTATGGCATGCACTGGGTCCGCCA
 GGCTCCAGGCAAGGGGCTGGAGTGGGTGGCATTATGACATATGATGGAAGTAATAGA
 TACTATGCAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATTCCAAGAACAC
 GCTGTATCTGCAAATGAACAGCCTGAGAGATGAGGACACGGCTCTATATTACTGTGCGA
 GAGATCGTATAGTGGGAGGTAGGGTCCCTGATGCTTTTGATATCTGGGGCCAAGGGAC
 AATGGTCACCGTCTCTTCA

SEQ ID NO. 452 (encodes V_H domain 3.2)

GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
 CTCTCCTGTGCAGCCTCTGGATTCCCCTTAATTAGCTATGGCATGAATTGGGTCCGCCA
 20 GGCTCCAGGCAAGGGGCTGGACTGGGTGGCATTATATCATATGATGGAAGTAATAAAT
 ATTATGCAGACTCCGTGAAGGGCCGATTACCATCTCCAAAGACAATTCCAAGAACACG
 CTGTATCTGCAAATGAACAGCCTGAGAGCTGAGGACACGGCTGTGTATTACTGTGCGAA
 AGATCGTATAGTGGGAGCCAGGGTCCCTGATGCTTTTGATATCTGGGGCCAAGGGACA
 ATGGTCACCGTCTCCTCA

SEQ ID NO. 453 (encodes V_H domain 3.3)

GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
 CTCTCCTGTGCAGCCTCTGGATTCCCCCTCATTAGCTATGGCATGAACTGGGTCCGCCA
 GGCTCCAGGCAAGGGGCTGGAGTGGGTGGCATTATATCATATGATGGAAGTAATAGAT
 ACTATGCAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATTCCAAGAACACG
 30 CTGTATCTGCAAATGAACAGCCTGAGAGCTGAGGACACGGCTGTATATTACTGTGCGAA
 AGATCGTATAGTGGGAGCTAGGGTCCCTGATGCTTTTGATATCTGGGGCCAAGGGACA
 ATGGTCACCGTCTCCTCA

SEQ ID NO. 454 (encodes V_H domain 3.4)

GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGCGGTCCAGCCTGGGAGGTCCCTGAG
ACTCTCCTGTGCAGCCTCTGGATTCCCCTTAATTAGCTATGGCATGAATTGGGTCCGCC
AGGCTCCAGGCAAGGGGCTGGACTGGGTGGCATTATAACATATGATGGAAGTAATAG
ATATTATGCAGACTCTGTGAAGGGCCGATTCACCATCTCCAGAGACAATTCCAAGAACA
CGCTTTATCTGCAAATGAACAGCCTGAGACCTGAGGACACGGCTGTATATTACTGTGCG
AAAGATCGTATTGTGGGAGCTAGGGTCCCTGATGCTTATGATATCTGGGGCCAAGGGA
CAATGGTCACCGTCTCCTCA

SEQ ID NO. 455 (encodes V_H domain 3.5)

10 CAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCCCCTTAATTAGCTATGGCATGAATTGGGTCCGCCA
GGCTCCAGGCAAGGGGCTGGACTGGGTGGCATTATAACATATGATGGAAGTAATAGAT
ACTATGCAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAATTCCAAGAACACG
CTTTATCTGCAAATGAACAGCCTGAGACCTGAGGACACGGCTGTATATTACTGTGCGAA
AGATCGTATTGTGGGAGCTAGGGTCCCTGATGCTTATGATATCTGGGGCCAAGGGACA
ATGGTCACCGTCTCCTCA

SEQ ID NO. 456 (encodes V_H domain 3.6)

20 GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCCCCTTAATTAGCTATGGCATGAATTGGGTCCGCCA
GGCTCCAGGCAAGGGGCTGGACTGGGTGGCATTATAACATATGATGGAAGTAATAGAT
ACTATGCAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAATTCCAAGAACACG
CTTTATCTGCAAATGAACAGCCTGAGACCTGAGGACACGGCTGTATATTACTGTGCGAA
AGATCGTATTGTGGGAGCTAGGGTCCCTGATGCTTATGATATCTGGGGCCAAGGGACA
ATGGTCACCGTCTCCTCA

SEQ ID NO. 457 (encodes V_H domain 3.7)

30 CAGGTGCAGCTGGTGGAGTCTGGGGGAGGCTTGGTACAGCCTGGGGGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCCCCTTAATTAGCTATGGCATGAATTGGGTCCGCCA
GGCTCCAGGCAAGGGGCTGGACTGGGTGGCATTATAACATATGATGGAAGTAATAGAT
ACTATGCAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAATTCCAAGAACACG
CTTCATCTGCAAATGGACAGCCTGAGACCTGAGGACACGGCTGTATATTACTGTGCGAA
AGATCGTATTGTGGGAGCTAGGGTCCCTGATGCTTATGATATCTGGGGCCAAGGGACA
ATGGTCACTGTCTCTTCA

SEQ ID NO. 458 (encodes V_H domain 3.8)

GAAGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCCCCTTAATTAGCTATGGCATGAATTGGGTCCGCCA

GGCTCCAGGCAAGGGGCTGGACTGGGTGGCATTATAACATATGATGGAAGTAATAGAT
ACTATGCAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAATCCAAGAACACG
CTTTATCTGCAAATGAACAGCCTGAGACCTGAGGACACGGCTGTATATTACTGTGCGAA
AGATCGTATTGTGGGAGCTAGGGTCCCTGATGCTTATGATATCTGGGGCCAGGGAACC
CTGGTCACTGTCTCCTCA

SEQ ID NO. 459 (encodes V_H domain 3.9)

CAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCCCCTTAATTAGCTATGGCATGAATTGGGTCCGCCA
GGCTCCAGGCAAGGGGCTGGAGTGGGTGGCATTATATCATATGATGGAAGTAATAGAT
10 ACTATGCAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAATCCAAGAACACG
CTGTATCTGCAAATGAACAGCCTGAGACCTGAGGACACGGCTGTATATTACTGTGCGAA
AGATCGTATTGTGGGAGCTAGGGTCCCTGATGCTTATGATATCTGGGGCCAAGGGACA
ATGGTCACCGTCTCTTCA

SEQ ID NO. 460 (encodes V_H domain 3.10)

CAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCCCCTTAATTAGCTATGGCATGAATTGGGTCCGCCA
GGCTCCAGGCAAGGGGCTGGACTGGGTGGCATTATAACATATGATGGAAGTAATAGAT
ACTATGCAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAATCCAAGAACACG
CTTCATCTGCAAATGAACAGCCTGAGACCTGAGGACACGGCTGTATATTACTGTGCGAA
20 AGATCGTATTGTGGGAGCTAGGGTCCCTGATGCTTATGATATCTGGGGCCAAGGGACA
ATGGTCACTGTCTCCTCA

SEQ ID NO. 461 (encodes V_H domain 3.11)

GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCCCCTTAATTAGCTATGGCATGAATTGGGTCCGCCA
GGCTCCAGGCAAGGGGCTGGACTGGGTGGCATTATAACATATGATGGAAGTAATAGAT
ACTATGCAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAATCCAAGAACACG
CTTCATCTGCAAATGAACAGCCTGAGACCTGAGGACACGGCTGTATATTACTGTGCGAA
AGATCGTATTGTGGGAGCTAGGGTCCCTGATGCTTATGATATCTGGGGCCAAGGGACA
ATGGTCACTGTCTCCTCA

30 SEQ ID NO. 462 (encodes V_H domain 3.12)

GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCCCCTTAATTAGCTATGGCATGAATTGGGTCCGCCA
GGCTCCAGGCAAGGGGCTGGACTGGGTGGCATTATAACATATGATGGAAGTAATAGAT
ACTATGCAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAATCCAAGAACACG

CTTTATCTGCAAATGAACAGCCTGAGACCTGAGGACACGGCTGTATATTACTGTGCGAA
AGATCGTATTGTGGGAGCTAGGGTCCCTGATGCTTATGATATCTGGGGCCAAGGGACA
ATGGTCACCGTCTCCTCA

SEQ ID NO. 463 (encodes V_H domain 3.13)

CAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCCCCTTAATTAGCTATGGCATGAATTGGGTCCGCCA
GGCTCCAGGCAAGGGGCTGGACTGGGTGGCATTATAACATATGATGGAAGTAATAGAT
ACTATGCAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAATTCCAAGAACACG
CTTTATCTGCAAATGAACAGCCTGAGACCTGAGGACACGGCTGTATATTACTGTGCGAA
10 AGATCGTATTGTGGGAGCTAGGGTCCCTGATGCTTATGATATCTGGGGCCAAGGGACA
ATGGTCACTGTCTCCTCA

SEQ ID NO. 464 (encodes V_H domain 3.14)

GAGGTGCAGCTGGTGGAGTCTGGGGGAGGTGTGGTACGGCCTGGGGGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCCCCTTAATTAGCTATGGCATGAATTGGGTCCGCCA
GGCTCCAGGCAAGGGGCTGGACTGGGTGGCATTATAACATATGATGGAAGTAATAGAT
ACTATGCAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAATTCCAAGAACACG
CTTCATCTGCAAATGAACAGCCTGAGACCTGAGGACACGGCTGTATATTACTGTGCGAA
AGATCGTATTGTGGGAGCTAGGGTCCCTGATGCTTATGATATCTGGGGCCAAGGGACA
ATGGTCACTGTCTCCTCA

20 SEQ ID NO. 465 (encodes V_H domain 3.15)

GAAGTGCAGCTGGTGGAGTCTGGGGGAGGCTTGGTCCAGCCTGGGGGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCCCCTTAATTAGCTATGGCATGAATTGGGTCCGCCA
GGCTCCAGGCAAGGGGCTGGACTGGGTGGCATTATAACATATGATGGAAGTAATAGAT
ACTATGCAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAATTCCAAGAACACG
CTTCATCTGCAAATGAACAGCCTGAGACCTGAGGACACGGCTGTATATTACTGTGCGAA
AGATCGTATTGTGGGAGCTAGGGTCCCTGATGCTTATGATATCTGGGGCCAAGGGACA
ATGGTCACCGTCTCCTCA

SEQ ID NO. 466 (encodes V_H domain 3.16)

GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
30 CTCTCCTGTGCAGCCTCTGGATTCCCCTTAATTAGCTATGGCATGAATTGGGTCCGCCA
GGCTCCAGGCAAGGGGCTGGACTGGGTGGCATTATAACATATGATGGAAGTAATAGAT
ACTATGCAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAATTCCAAGAACACG
CTTCATCTGCAAATGAACAGCCTGAGACCTGAGGACACGGCTGTATATTACTGTGCGAA

AGATCGTATTGTGGGAGCTAGGGTCCCTGATGCTTATGATATCTGGGGCCAAGGGACA
ATGGTCACCGTCTCCTCA

SEQ ID NO. 467 (encodes V_H domain 3.17)

GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCCCCTTAATTAGCTATGGCATGAATTGGGTCCGCCA
GGCTCCAGGCAAGGGGCTGGACTGGGTGGCATTATAACATATGATGGAAGTAATAGAT
ACTATGCAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAATTCCAAGAACACG
CTTTATCTGCAAATGAACAGCCTGAAACCTGAGGACACGGCTGTATATTACTGTGCGAA
AGATCGTATTGTGGGAGCTAGGGTCCCTGATGCTTATGATATCTGGGGCCAAGGGACA
10 ATGGTCACCGTCTCCTCA

SEQ ID NO. 468 (encodes V_H domain 3.18)

GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCCCCTTAATTAGCTATGGCATGAATTGGGTCCGCCA
GGCTCCAGGCAAGGGGCTGGACTGGGTGGCATTATAACATATGATGGAAGTAATAGAT
ACTATGCAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAATTCCAAGAACACG
CTTTATCTGCAAATGAACAGCCTGAAACCTGAGGACACGGCTGTATATTACTGTGCGAA
AGATCGTATTGTGGGAGCTAGGGTCCCTGATGCTTATGATATCTGGGGCCAAGGGACA
ATGGTCACCGTCTCCTCA

SEQ ID NO. 469 (encodes V_H domain 3.19)

20 CAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCCCCTTAATTAGCTATGGCATGCACTGGGTCCGCCA
GGCTCCAGGCAAGGGGCTGGAGTGGGTGGCATTATGACATATGATGGAAGTAATAGA
TACTATGCAGACGCCGTGAAGGGCCGATTCACCATCTCCAGAGACAATTCCAAGAACAC
GCTGTATCTGCAAATGAACAGCCTGAGAGCTGAGGACACGGCTGTGTATTACTGTGCGA
GAGATCGTATAGTGGGAGGTAGGGTCCCTGATGCTTTTGATATCTGGGGCCAAGGGAC
AATGGTCACCGTCTCTTCA

SEQ ID NO. 470 (encodes V_H domain 3.20)

CAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCCCCTTAATTAGCTATGGCATGCACTGGGTCCGCCA
30 GGCTCCAGGCAAGGGGCTGGAGTGGGTGGCATTTCAGACATATGATGGCAGTAATAGA
TACTATGCAGACGCCGTGAAGGGCCGATTCACCATCTCCAGAGACAATTCCAAGAACAC
GCTGTATCTGCAAATGAACAGCCTGAGAGCTGAGGACACGGCTGTGTATTACTGTGCGA
GAGATCGTATAGTGGGAGGTAGGGTCCCTGATGCTTTTGATATCTGGGGCCAAGGGAC
AATGGTCACCGTCTCTTCA

SEQ ID NO. 471 (encodes V_H domain 3.21)

CAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCCCCTTAATTAGCTATGGCATGCACTGGGTCCGCCA
GGCTCCAGGCAAGGGGCTGGAGTGGGTGGCATTTCAGACATATGATGGCAGTAATAGA
TACTATGCAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATTCCAAGAACAC
GCTGTATCTGCAAATGAACAGCCTGAGAGCTGAGGACACGGCTGTGTATTACTGTGCGA
GAGATCGTATAGTGGGAGGTAGGGTCCCTGATGCTTTTGATATCTGGGGCCAAGGGAC
AATGGTCACCGTCTCTTCA

SEQ ID NO. 472 (encodes V_H domain 3.22)

10 CAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCCCCTTAATTAGCTATGGCATGCACTGGGTCCGCCA
GGCTCCAGGCAAGGGGCTGGAGTGGGTGGCATTTCAGACATATGATGCCAGTAATAGA
TACTATGCAGACGCCGTGAAGGGCCGATTACCATCTCCAGAGACAATTCCAAGAACAC
GCTGTATCTGCAAATGAACAGCCTGAGAGCTGAGGACACGGCTGTGTATTACTGTGCGA
GAGATCGTATAGTGGGAGGTAGGGTCCCTGATGCTTTTGATATCTGGGGCCAAGGGAC
AATGGTCACCGTCTCTTCA

SEQ ID NO. 473 (encodes V_H domain 3.23)

CAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCCCCTTAATTAGCTATGGCATGAATTGGGTCCGCCA
20 GGCTCCAGGCAAGGGGCTGGAGTGGGTGGCATTATAACATATGATGGAAGTAATAGA
TACTATGCAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATTCCAAGAACAC
GCTTTATCTGCAAATGAACAGCCTGAGAGCTGAGGACACGGCTGTATATTACTGTGCGA
AAGATCGTATTGTGGGAGCTAGGGTCCCTGATGCTTATGATATCTGGGGCCAAGGGAC
AATGGTCACTGTCTCCTCA

SEQ ID NO. 474 (encodes V_H domain 3.24)

AGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGAC
TCTCCTGTGCAGCCTCTGGATTCCCCTTAATTAGCTATGGCATGAATTGGGTCCGCCAG
GCTCCAGGCAAGGGGCTGGAGTGGGTGGCATTATAACATATGATGGAAGTAATAGATA
CTATGCAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATTCCAAGAACACGC
30 TTTATCTGCAAATGAACAGCCTGAGAGCTGAGGACACGGCTGTATATTACTGTGCGAAA
GATCGTATTGTGGGAGCTAGGGTCCCTGATGCTTATGATATCTGGGGCCAAGGGACAAT
GGTCACTGTCTCCTCA

family 4

SEQ ID NO. 475 (encodes V_H domain 4.1)

CAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCTTGAGA
CTCTCCTGTGTAGCCTCTGGATTCCCCTTCATTAGCTATGGCATGCACTGGGTCCGCCA
GGCTCCAGGCAAGGGGCGGGAGTGGGTGGCAGTTATATCATATGATGGAAGTAATAGA
TACTATGCAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATTCCAAGAACAC
GCTGTATCTGCAAATGAACAGCCTGAGACCTGAGGACACGGCTGTGTATTATTGTGCGA
AAGAGAGGATTTTTGGAGTGCTTACCCCTGATGATTTTGATATCTGGGGCCAAGGGACA
ACGGTCACCGTCTCCTCA

SEQ ID NO. 476 (encodes V_H domain 4.2)

10 CAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCTTGAGA
CTCTCCTGTGCAGCCTCTGGATTCCCCTTCATTAGCTATGGCATGCACTGGGTCCGCCA
GGCTCCAGGCAAGGGGCTGGAGTGGGTGGCAGTTATATCATATGATGGAAGTAATAGA
TACTATGCAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATTCCAAGAACAC
GCTGTATCTGCAAATGAACAGCCTGAGACCTGAGGACACGGCTGTGTATTACTGTGCGA
AAGAGAGGATTTTTGGAGTGCTTACCCCTGATGATTTTGATATCTGGGGCCAAGGGACA
ACGGTCACTGTCTCCTCA

SEQ ID NO. 477 (encodes V_H domain 4.3)

GAGGTGCAGCTGTTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCTTGAGA
CTCTCCTGTGCAGCCTCTGGATTCCCCTTCATTAGCTATGGCATGCACTGGGTCCGCCA
GGCTCCAGGCAAGGGGCTGGAGTGGGTGGCAGTTATATCATATGATGGAGCTAATAGA
20 TACTATGCAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATTCCAAGAACAC
GCTGTATCTGCAAATGAACAGCCTGAGACCTGAGGACACGGCTGTGTATTATTGTGCGA
AAGAGAGGATTTTTGGCGTGCTTACCCCTGATGATTTTGAAATCTGGGGCCAAGGGACA
ACGGTCACCGTCTCCTCA

SEQ ID NO. 478 (encodes V_H domain 4.4)

GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCTTGAGA
CTCTCCTGTGCAGCCTCTGGATTCACCTTCACTAGCTATGGCATGCACTGGGTCCGCCA
GGCTCCAGGCAAGGGGCTGGAGTGGGTGGCAGTTATATCATATGATGGAAGTAATAGA
TACTATGCAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATTCCAAGAACAC
GCTGTATCTGCAAATGAACAGCCTGAGACCTGAGGACACGGCTGTGTATTACTGTGCGA
30 AAGAGAGGATTTTTGGAGCGCTTACCCCTGATGATTTTGATATCTGGGGCCAAGGGACA
ACGGTCACCGTCTCTTCA

family 5

SEQ ID NO. 479 (encodes V_H domain 5.1)

CAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTACCTTCAATAACTATGGCATGCACTGGGTCCGCCA
GGCTCCAGGCAAGGGGCTGGAGTGGGTGGCAATTATATCATATGATGGAAATACTAAAT
ATTATACAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATTCCAAGAACACG
CTGTATCTGCAAATGAATAGCCTGAGAGTTGAGGACACGGCTGTGTATTACTGTGCGAA
AGGTTTATGGCCTTCGGACGTCTGGGGCCAAGGGACCACGGTCACTGTCTCTTCA

SEQ ID NO. 480 (encodes V_H domain 5.2)

GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTACCTTCAATAACTATGGCATGCACTGGGTCCGCCA
10 GGCTCCAGGCAAGGGGCTGGAGTGGGTGGCAATTATATCATATGATGGAAATAGTAAAT
ATTATACAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATTCCAAGAACACG
CTGTATCTGCAAATGAATAGCCTGAGAGTTGAGGACACGGCTGTGTATTACTGTGCGAA
AGGTTTATGGCCTTCGGACGTCTGGGGCCAAGGGACCACGGTCACCGTCTCCTCA

family 6

SEQ ID NO. 481 (encodes V_H domain 6.1)

CAGGTGCAGCTACAGGAGTCGGGCCCAGGACTGGTGAAGCCTTCACAGACCCTGTCCC
TCACCTGCACTGTCTCTGGTGGCTCCATCAGCAATAGTGGTTATTACTGGAGCTGGGTC
CGCCAGCACCCAGGGAAGGACCTGGAGTGGATTGGGTTTCATCTATTACAATGGGAGCA
TCCACTACAACCCGTCCCTCAAGAGTCGAGTTATCATATCAGTAGACACGTCTAAGAAC
20 CAGTTCTCCCTGAAAATGAACTCTGTGACTGCCGCGGACACGGCCGTGTATTACTGTGC
GAGAGACGGGGATGACTACGGTGACTACTTGAGGGGGCCAGGGAACCCTGGTCACCGT
CTCCTCA

SEQ ID NO. 482 (encodes V_H domain 6.2)

CAGGTGCAGCTGCAGGAGTCGGGCCCAGGACTGGTGAAGCCTTCACAGACCCTGTCC
CTCACCTGCACTGTCTCTGGTGGCTCCATCAGCAATAGTGGTTATTACTGGAGCTGGAT
CCGCCAGCACCCAGGGAAGGGCCTGGAGTGGATTGGGTTTCATCTATTACAATGGGAGC
ATCCACTACAACCCGTCCCTCAAGAGTCGAGTTATCATATCAGTAGACACGTCTAAGAA
CCAGTTCTCCCTGAAAATGAGCTCTGTGACTGCCGCGGACACGGCCGTGTATTACTGTG
CGAGAGACGGGGATGACTACGGTGACTACTTGAGGGGGCCAGGGAACCCTGGTCACCG
30 TCTCCTCA

SEQ ID NO. 483 (encodes V_H domain 6.3)

CAGGTGCAGCTGCAGGAGTCGGGCCCAGGACTGGTGAAGCCTTCACAGACCCTGTCC
CTCACCTGCACTGTCTCTGGTGGCTCCATCAGCAATAGTGGTTATTACTGGAGCTGGGT
CCGCCAGCACCCAGGGAAGGGCCTGGAGTGGATTGGGTTTCATCTATTACAATGGGAGC

ATCCACTACAACCCGTCCCTCAAGAGTCGAGTTATCATATCAGTAGACACGTCTAAGAA
CCAGTTCTCCCTGAAACTGAACTCTGTGACTGCCGCGGACACGGCCGTGTATTACTGTG
CGAGAGACGGGGATGACTACGGTGACTACTTGAGGGGCCAGGGAACCCTGGTCACCG
TCTCCTCA

SEQ ID NO. 484 (encodes V_H domain 6.4)

CAGGTGCAGCTGCAGGAGTCGGGCCCAGGACTGGTGAAGCCTTCACAGACCCTGTCC
CTCACCTGCACTGTCTCTGGTGGCTCCATCAGCAATAGTGGTTATTACTGGAGCTGGAT
CCGCCAGCACCCAGGGAAGGGCCTGGAGTGGATTGGGTTCATCTATTACAATGGGAGC
ATCCACTACAACCCGTCCCTCAAGAGTCGAGTTATCATATCAGTAGACACGTCTAAGAA
10 CCAGTTCTCCCTGAAACTGAGCTCTGTGACTGCCGCGGACACGGCCGTGTATTACTGT
GCGAGAGACGGGGATGACTACGGTGACTACTTGAGGGGCCAGGGAACCCTGGTCACC
GTCTCCTCA

SEQ ID NO. 485 (encodes V_H domain 6.5)

CAGGTGCAGCTGCAGGAGTCGGGCCCAGGACTGGTGAAGCCTTCACAGACCCTGTCC
CTCACCTGCACTGTCTCTGGTGGCTCCATCAGCAATAGTGGTTATTACTGGAGCTGGGT
CCGCCAGCACCCAGGGAAGGGCCTGGAGTGGATTGGGTTCATCTATTACAATGGGAGC
ATCCACTACAACCCGTCCCTCAAGAGTCGAGTTACCATATCAGTAGACACGTCTAAGAA
CCAGTTCTCCCTGAAAATGAGCTCTGTGACTGCCGCGGACACGGCCGTGTATTACTGTG
CGAGAGACGGGGATGACTACGGTGACTACTTGAGGGGCCAGGGAACCCTGGTCACCG
20 TCTCCTCA

SEQ ID NO. 486 (encodes V_H domain 6.6)

CAGGTGCAGCTGCAGGAGTCGGGCCCAGGACTGGTGAAGCCTTCACAGACCCTGTCC
CTCACCTGCACTGTCTCTGGTGGCTCCATCAGCAATAGTGGTTATTACTGGAGCTGGGT
CCGCCAGCACCCAGGGAAGGGCCTGGAGTGGATTGGGTTCATCTATTACAATGGGAGC
ATCCACTACAACCCGTCCCTCAAGAGTCGAGTTACCATATCAGTAGACACGTCTAAGAA
CCAGTTCTCCCTGAAACTGAACTCTGTGACTGCCGCGGACACGGCCGTGTATTACTGTG
CGAGAGACGGGGATGACTACGGTGACTACTTGAGGGGCCAGGGAACCCTGGTCACCG
TCTCCTCA

SEQ ID NO. 487 (encodes V_H domain 6.7)

30 CAGGTGCAGCTGCAGGAGTCGGGCCCAGGACTGGTGAAGCCTTCACAGACCCTGTCC
CTCACCTGCACTGTCTCTGGTGGCTCCATCAGCAATAGTGGTTATTACTGGAGCTGGGT
CCGCCAGCACCCAGGGAAGGGCCTGGAGTGGATTGGGTTCATCTATTACAATGGGAGC
ATCCACTACAACCCGTCCCTCAAGAGTCGAGTTACCATATCAGTAGACACGTCTAAGAA
CCAGTTCTCCCTGAAACTGAGCTCTGTGACTGCCGCGGACACGGCCGTGTATTACTGT

GCGAGAGACGGGGATGACTACGGTGACTACTTGAGGGGCCAGGGAACCCTGGTCACC
GTCTCCTCA

family 7

SEQ ID NO. 488 (encodes V_H domain 7.1)

GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCTTGGTCCAGCCTGGGGGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCACCTTTAGTAGCTATTGGATGTACTGGGTCCGCCA
GGCTCCAGGGAAGGGGCTGGAGTGGGTGGCCAACATAAATCACGATGGAAGTGAGAA
ATACTATGTGGA CTCTGTGAAGGGCCGATTCACCATCTCCAGAGACAACGCCAAGA AACT
CACTGTATCTGCAAATGAACAGCCTGAGAGCCGAGGACACGGCTGTGTATTACTGCGC
10 GAGAGATTCCCTTATAGTGGGAGAGAGGGGCTACTGGGGCCAGGGAACCCTGGTCAC
CGTCTCCTCA

SEQ ID NO. 489 (encodes V_H domain 7.2)

GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCTTGGTCCAGCCTGGGGGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCACCTTTAGTAGCTATTGGATGTACTGGGTCCGCCA
GGCTCCAGGGAAGGGGCTGGAGTGGGTGGCCAACATAAATCACGATGGAAGTGAGAA
ATACTATGTGGA CTCTGTGAAGGGCCGATTCACCATCTCCAGAGACAACGCCAAGA AACT
CACTGTATCTGCAAATGAACAGCCTGAGAGCCGAGGACACGGCTGTGTATTACTGCGC
GAGAGATAACCTTATAGTGGGAGAGAGGGGCTACTGGGGCCAGGGAACCCTGGTCAC
CGTCTCCTCA

20 SEQ ID NO. 490 (encodes V_H domain 7.3)

GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCTTGGTCCAGCCTGGGGGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCACCTTTAGTAGCTATTGGATGTACTGGGTCCGCCA
GGCTCCAGGGAAGGGGCTGGAGTGGGTGGCCAACATAAATCACGGGGGAAGTGAGAA
ATACTATGTGGA CTCTGTGAAGGGCCGATTCACCATCTCCAGAGACAACGCCAAGA AACT
CACTGTATCTGCAAATGAACAGCCTGAGAGCCGAGGACACGGCTGTGTATTACTGCGC
GAGAGATTCCCTTATAGTGGGAGAGAGGGGCTACT

SEQ ID NO. 491 (encodes V_H domain 7.4)

GGGGCCAGGGAACCCTGGTCACCGTCTCCTCA

GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCTTGGTCCAGCCTGGGGGGTCCCTGAGA
30 CTCTCCTGTGCAGCCTCTGGATTCACCTTTAGTAGCTATTGGATGTACTGGGTCCGCCA
GGCTCCAGGGAAGGGGCTGGAGTGGGTGGCCAACATAAATCACCAGGGAAGTGAGAA
ATACTATGTGGA CTCTGTGAAGGGCCGATTCACCATCTCCAGAGACAACGCCAAGA AACT
CACTGTATCTGCAAATGAACAGCCTGAGAGCCGAGGACACGGCTGTGTATTACTGCGC

GAGAGATTCCCTTATAGTGGGAGAGAGGGGCTACTGGGGCCAGGGAACCCTGGTCAC
CGTCTCCTCA

SEQ ID NO. 492 (encodes V_H domain 7.5)

GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCTTGGTCCAGCCTGGGGGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTACCTTTAGTAGCTATTGGATGTACTGGGTCCGCCA
GGCTCCAGGGAAGGGGCTGGAGTGGGTGGCCAACATAAATCACCCCGGAAGTGAGAA
ATACTATGTGGACTCTGTGAAGGGCCGATTCACCATCTCCAGAGACAACGCCAAGAAGT
CACTGTATCTGCAAATGAACAGCCTGAGAGCCGAGGACACGGCTGTGTATTACTGCGC
GAGAGATTCCCTTATAGTGGGAGAGAGGGGCTACTGGGGCCAGGGAACCCTGGTCAC
10 CGTCTCCTCA

SEQ ID NO. 493 (encodes V_H domain 7.6)

GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCTTGGTCCAGCCTGGGGGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTACCTTTAGTAGCTATTGGATGTACTGGGTCCGCCA
GGCTCCAGGGAAGGGGCTGGAGTGGGTGGCCAACATAAATCACGAGGGAAGTGAGAA
ATACTATGTGGACTCTGTGAAGGGCCGATTCACCATCTCCAGAGACAACGCCAAGAAGT
CACTGTATCTGCAAATGAACAGCCTGAGAGCCGAGGACACGGCTGTGTATTACTGCGC
GAGAGATTCCCTTATAGTGGGAGAGAGGGGCTACTGGGGCCAGGGAACCCTGGTCAC
CGTCTCCTCA

SEQ ID NO. 494 (encodes V_H domain 7.7)

20 GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCTTGGTCCAGCCTGGGGGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTACCTTTAGTAGCTATTGGATGTACTGGGTCCGCCA
GGCTCCAGGGAAGGGGCTGGAGTGGGTGGCCAACATAAATCACATCGGAAGTGAGAAA
TACTATGTGGACTCTGTGAAGGGCCGATTCACCATCTCCAGAGACAACGCCAAGAAGT
ACTGTATCTGCAAATGAACAGCCTGAGAGCCGAGGACACGGCTGTGTATTACTGCGCG
AGAGATTCCCTTATAGTGGGAGAGAGGGGCTACTGGGGCCAGGGAACCCTGGTCACC
GTCTCCTCA

SEQ ID NO. 495 (encodes V_H domain 7.8)

GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCTTGGTCCAGCCTGGGGGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTACCTTTAGTAGCTATTGGATGTACTGGGTCCGCCA
30 GGCTCCAGGGAAGGGGCTGGAGTGGGTGGCCAACATAAATCACGATGGAAGTGAGAA
ATACTATGTGGACTCTGTGAAGGGCCGATTCACCATCTCCAGAGACAACGCCAAGAAGT
CACTGTATCTGCAAATGAACAGCCTGAGAGCCGAGGACACGGCTGTGTATTACTGCGC
GAGAGATACCCTTATAGTGGGAGAGAGGGGCTACTGGGGCCAGGGAACCCTGGTCAC
CGTCTCCTCA

family 8

SEQ ID NO. 496

CAGGTGCAGCTACAGCAGTGGGGCGCAGGACTGTTGAAGCCTTCGGAGACCCTGTCC
CTCACCTGCGCTGTCTATGGTGGGTCCTTCAGTGGTTACTACTGGAGCTGGATCCGCCA
GCCCCAGGGAAGGGGCTGGAGTGGATTGGGGAAATCAATCATAGTGGAAGCACCAAC
TACAACCCGTCCCTCAAGAGTCGAGTCACCATATCAGTAGACACGTCCAAGAACCAGTT
CTCCCTGAAGCTGAGCTCTGTGACCGCCGCGGACACGGCTGTGTATTACTGTGCGAGA
GGCCCCATACCAGCCACTGCTATACCCGATGCTTTTGATATCTGGGGCCAAGGGACAAT
GGTCACTGTCTCCTCA

10 family 9

SEQ ID NO. 497

GAGGTGCAGCTACAGCAGTGGGGCGCAGGACTGTTGAAGCCTTCGGAGACCCTGTCC
CTCACCTGCGCTGTCTATGGTGGGTCCTTCAGTGGTCACTACTGGAGCTGGATCCGCC
AGCCCCAGGGAAGGGGCTGGAGTGGATTGGGGACATAAATCATAGTGGAAGCACCA
CTACAACCCGTCCCTCAAGAGTCGAGTCACCATATCAGTAGACACGTCCAAGAATCAGT
TCTCCCTGAAGCTGAGCTCTGTGACCGCCGCGGACACGGCTGTGTATTACTGTGTGAG
AGACTACGGTGACTCCCGTAGCCTTTTTGACTACTGGGGCCAGGGAACCCTGGTCACC
GTCTCTTCA

family 10

20 SEQ ID NO. 498

CAGGTGCAGCTGGTGGAGTCTGGGGGAGGCTTGGTACAGCCTGGGGGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTACCTTCAGTAGCTATGGCATGCACTGGGTCCGCCA
GGCTCCAGGCAAGGGGCTGGAGTGGGTGGCATTATGTCATATGATGGCAGTAATAAA
TACTATGTAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATTCCAAGAATAC
GCTGTATCTGCAAATGAACAGCCTGAGAGCTGAGGACACGGCTGTGTATTACTGTGCGA
AAGGCGATTACGATTTTTGGAGTGGTTACCCCGACTACGATATGGACGTCTGGGGCCAA
GGGACCACGGTCACCGTCTCCTCA

family 11

SEQ ID NO. 499

30 GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCTTGGTCAAGCCTGGAGGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTCAACTTGATTAGCTATGGCATGTACTGGGTCCGCCA
GGCTCCAGGCAAGGGGCTGGAGTGGGTGGCAGTTATATCATATGATGGAAGTAATAAA
AACTATGCAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATTCCAAGAATAC

GCTGTTTCTGCAAATGAACAGCCTGAGAGTTGAGGACACGGCTGTGTATTACTGTGCGA
AAGGGGGGAATGCCTTGTATAGCAGTGGCTGGCCCGATGATGGTTTTGATATCAGGGG
CCAAGGGACAATGGTCACTGTCTCCTCA

family 12

SEQ ID NO. 500

CAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTACCTTCAGTAACTTTGGCATGCACTGGGCCCCGCCA
GGCTCCAGGCAAGGGACTGGAGTGGGTGGCAGTAATATCATATGATGGAAATAGTAAAT
ACTATGCAGACACCGTGAAGGGCCGATTCACCATCTCCAGAGACAATTCCAAGAACACG
10 CTGTATCTGGAAATGAACAGCCTGAGAGCTGATGACACGGCTGTGTATTACTGTGCGAA
AGGCCTATGGCCCCCAATGGACGTCAGGGGCCAAGGGACCACGGTCACCGTCTCCTC
A

family 13

SEQ ID NO. 501

GAGGTGCAGCTGGTGGAGTCTGGGGGAGGCTCGGTCCAGCCTGGGGGGTCCCTGAGA
CTCTCCTGTGCAGCCTCTGGATTACCTTTAGTGA CTATTGGATGACCTGGGTCCGCCA
GGTTCCAGGGAAGGGGCTGGAGTGGGTGGCCAACATAAAGCAAGATGGAAGTGAGAA
ATACTATGTGGACTCTGTGAAGGGCCGATTCACCATCTCCAGAGACAACGCCAAGAACT
CACTATATCTGCAAATGAATAGCCTGAGAGCCGAGGACACGGCTGTGTATTACTGTGCG
20 AGAGATCGAGGAGGAGCAGTGGCCCTTTATCACAACGGTATGGACATGGGGGGCCAAG
GGACCACGGTCACTGTCTCTTCA

family 14

SEQ ID NO. 502

GAAGTGCAGCTGGTGGAGTCTGGGGGAGGTGAAGAAGCCTGGGGCCTCAGTGAAGGT
CTCCTGCAAGGCTTCTGGATACACCTTCACCA GTTATGATATCAACTGGGTGCGACAGG
CCACTGGACAAGGGCTTGAGTGGATGGGATGGATGAACCCTAACAGTGGTAACACAGG
CTATGCACAGAAGTTCCAGGGCAGAGTCACCATGACCAGGAACACCTCCATAAGCACA
GCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGACACGGCCGTGTATTACTGTGCGA
GAGGGAACGGGCCCCGGTATAACTGGA ACTACTGACTACTGGGGCCAGGGAACCCTGG
30 TCACTGTCTCTTCA

family 15

SEQ ID NO. 503

GAGGTGCAGCTGGTGGAGTCTGGGGGAGGTGTGGTACGGCCTGGGGGGTCCCTGAGA
 CTCTCCTGTGCAGCCTCTGGATTACCTTTGATGATTATGGCATGAGCTGGGTCCGCCA
 AGCTCCAGGGAAGGGGCTGGAGTGGGTCTCTGGTATTAATTGGAATGGTGATCGTACC
 GGTTATGCAGACTCTGTGAAGGGCCGATTACCATCTCCAGAGACAACGCCAAGAAGCTC
 CCTGTATCTGCAAATGAACAGTCTGAGAGCCGAGGACACGGCCTTGTATTACTGTGGGA
 GAGAGAATGTTATAGTACCAGCTGCTACCTACTGGGGCCAGGGAACCCTGGTCACCGT
 CTCCTCA

10 The term "nucleic acid," "polynucleotide," or "nucleic acid molecule" refers to deoxyribonucleic acid (DNA) or ribonucleic acid (RNA), or a combination of a DNA or RNA. RNA includes *in vitro* transcribed RNA, synthetic RNA or mRNA sequence. The nucleic acid construct may further comprise a suicide gene. The construct may be in the form of a plasmid, vector, transcription or expression cassette.

In another aspect, the invention relates to an isolated nucleic acid construct comprising a nucleic acid as defined above. The construct may be in the form of a plasmid, vector, transcription or expression cassette.

20 The nucleic acid can be cloned into a number of types of vectors. For example, the nucleic acid can be cloned into a vector including, but not limited to a plasmid, a phagemid, a phage derivative and a cosmid. Vectors of particular interest include expression vectors, replication vectors, probe generation vectors, and sequencing vectors. The vectors can be suitable for replication and integration eukaryotes. Typical cloning vectors contain transcription and translation terminators, initiation sequences, and promoters useful for regulation of the expression of the desired nucleic acid sequence.

In one embodiment, the vector is an *in vitro* transcribed vector, e.g., a vector that transcribes RNA of a nucleic acid molecule described herein. The expression vector may be provided to a cell in the form of a viral vector. Viral vector technology is well known in the art and is described, for example, in Sambrook et al. (Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, New York, 2013). A number of viral based systems have been developed for gene transfer into mammalian cells. For example, retroviruses such as adenovirus vectors can be used. In one embodiment, a lentivirus vector is used.

30 In a further aspect, the invention also relates to an isolated cell or cell population comprising one or more nucleic acid construct or vector as described above. In one embodiment, the cell is an isolated recombinant host cell comprising one or more nucleic acid construct as described above. The host cell may be a bacterial, viral, plant, mammalian or other suitable host cell.

Such host cells are well known in the art and many are available from the American Type Culture Collection (ATCC). These host cells include, inter alia, Chinese hamster ovary (CHO) cells, NSO, SP2 cells, HeLa cells, baby hamster kidney (BHK) cells, monkey kidney cells (COS), human hepatocellular carcinoma cells (e.g., Hep G2), A549 cells, 3T3 cells, HEK-293 cells and a number of other cell lines. Mammalian host cells include human, mouse, rat, dog, monkey, pig, goat, bovine, horse and hamster cells. Other cell lines that may be used are insect cell lines (e.g., *Spodoptera frugiperda* or *Trichoplusia ni*), amphibian cells, bacterial cells, plant cells and fungal cells. Fungal cells include yeast and filamentous fungus cells including, for example, *Pichia pastoris*, *Pichia finlandica*, *Pichia trehalophila*, *Pichia koclamae*,
 10 *Pichia membranaefaciens*, *Pichia minuta* (*Ogataea minuta*, *Pichia lindneri*), *Pichia opuntiae*, *Pichia thermotolerans*, *Pichia salictaria*, *Pichia guercuum*, *Pichia pijperi*, *Pichia stiptis*, *Pichia methanolica*, *Pichia* sp., *Saccharomyces cerevisiae*, *Saccharomyces* sp., *Hansenula polymorpha*, *Kluyveromyces* sp., *Kluyveromyces lactis*, *Candida albicans*, *Aspergillus nidulans*, *Aspergillus niger*, *Aspergillus oryzae*, *Trichoderma reesei*, *Chrysosporium lucknowense*, *Fusarium* sp., *Fusarium gramineum*, *Fusarium venenatum*, *Physcomitrella*.

Pharmaceutical compositions

In another aspect of the present invention, there is provided a pharmaceutical composition comprising single domain antibody or binding molecule according to the present invention and optionally a pharmaceutically acceptable carrier.

20 The genetically modified cells or pharmaceutical composition of the present invention can be administered by any convenient route, including parenteral administration. Parenteral administration includes, for example, intravenous, intramuscular, intraarterial, intraperitoneal, intranasal, rectal, intravesical, intradermal, topical or subcutaneous administration. Compositions can take the form of one or more dosage units.

The composition of the invention can be in the form of a liquid, e.g., a solution, emulsion or suspension. The liquid can be useful for delivery by injection, infusion (e.g., IV infusion) or subcutaneously. The liquid compositions of the invention, whether they are solutions, suspensions or other like form, can also include one or more of the following: sterile diluents such as water, saline solution, preferably physiological saline, Ringer's solution, isotonic sodium chloride,
 30 fixed oils such as synthetic mono or diglycerides, polyethylene glycols, glycerin, or other solvents; antibacterial agents such as benzyl alcohol or methyl paraben; and agents for the adjustment of tonicity such as sodium chloride or dextrose. A composition can be enclosed in an ampoule, a disposable syringe or a multiple-dose vial made of glass, plastic or other material.

The amount of the pharmaceutical composition of the present invention that is effective/active in the treatment of a particular disorder or condition will depend on the nature of the disorder or condition, and can be determined by standard clinical techniques. In addition, *in vitro* or *in vivo* assays can optionally be employed to help identify optimal dosage ranges. The precise dose to be employed in the compositions will also depend on the route of administration, and the seriousness of the disease or disorder, and should be decided according to the judgment of the practitioner and each patient's circumstances.

10 The compositions of the invention comprise an effective amount of a binding molecule of the present invention such that a suitable dosage will be obtained. The correct dosage of the compounds will vary according to the particular formulation, the mode of application, and its particular site, host and the disease being treated. Other factors like age, body weight, sex, diet, time of administration, rate of excretion, condition of the host, drug combinations, reaction sensitivities and severity of the disease shall be taken into account. Administration can be carried out continuously or periodically within the maximum tolerated dose.

Typically, this amount is at least about 0.01% of a binding molecule of the present invention by weight of the composition.

Preferred compositions of the present invention are prepared so that a parenteral dosage unit contains from about 0.01 % to about 2% by weight of the binding molecule of the present invention.

20 For intravenous administration, the composition can comprise from typically about 0.1 mg/kg to about 250 mg/kg of the animal's body weight, preferably, between about 0.1 mg/kg and about 20 mg/kg of the animal's body weight, and more preferably about 1 mg/kg to about 10 mg/kg of the animal's body weight.

The present compositions can take the form of suitable carriers, such as aerosols, sprays, suspensions, or any other form suitable for use. Other examples of suitable pharmaceutical carriers are described in "Remington's Pharmaceutical Sciences" by E. W. Martin.

30 The pharmaceutical compositions can be prepared using methodology well known in the pharmaceutical art. For example, a composition intended to be administered by injection can be prepared by combining a binding molecule of the present invention with water so as to form a solution. A surfactant can be added to facilitate the formation of a homogeneous solution or suspension.

The pharmaceutical composition of the invention can be co-administered with other therapeutics, for example anti-cancer agents.

Exemplary combinations with other agents

The molecules or pharmaceutical composition of the invention may be administered as the sole active ingredient or in combination with one or more other therapeutic agent. A therapeutic agent is a compound or molecule which is useful in the treatment of a disease. Examples of therapeutic agents include antibodies, antibody fragments, drugs, toxins, nucleases, hormones, immunomodulators, pro-apoptotic agents, anti-angiogenic agents, boron compounds, photoactive agents or dyes and radioisotopes. An antibody molecule includes a full antibody or fragment thereof (e.g., a Fab, F(ab')₂, Fv, a single chain Fv fragment (scFv) or a single domain antibody, for example a V_H domain, or antibody mimetic protein.

- 10 In one embodiment, the single variable heavy chain domain antibody that binds to MSLN, a binding molecule comprising a single variable heavy chain domain antibody that binds to MSLN or pharmaceutical composition described herein is used in combination with an existing therapy or therapeutic agent, for example an anti-cancer therapy. Thus, in another aspect, the invention also relates to a combination therapy comprising administration of a single variable heavy chain domain antibody that binds to MSLN, a binding molecule comprising a single variable heavy chain domain antibody that binds to MSLN or pharmaceutical composition described herein and an anti-cancer therapy.

- 20 The anti-cancer therapy may include a therapeutic agent or radiation therapy and includes gene therapy, viral therapy, RNA therapy bone marrow transplantation, nanotherapy, targeted anti-cancer therapies or oncolytic drugs. Examples of other therapeutic agents include other checkpoint inhibitors, antineoplastic agents, immunogenic agents, attenuated cancerous cells, tumor antigens, antigen presenting cells such as dendritic cells pulsed with tumor-derived antigen or nucleic acids, immune stimulating cytokines (e.g., IL-2, IFN α 2, GM-CSF), targeted small molecules and biological molecules (such as components of signal transduction pathways, e.g. modulators of tyrosine kinases and inhibitors of receptor tyrosine kinases, and agents that bind to tumor- specific antigens, including EGFR antagonists), an anti-inflammatory agent, a cytotoxic agent, a radiotoxic agent, or an immunosuppressive agent and cells transfected with a gene encoding an immune stimulating cytokine (e.g., GM-CSF), chemotherapy. In one embodiment, the single domain antibody is used in combination with
- 30 surgery.

In a specific embodiment of the present invention, the single variable heavy chain domain antibody that binds to MSLN, a binding molecule comprising a single variable heavy chain domain antibody that binds to MSLN or a pharmaceutical composition described herein is administered concurrently with a chemotherapeutic agent or with radiation therapy. In another specific embodiment, the chemotherapeutic agent or radiation therapy is administered prior or

subsequent to administration of the composition of the present invention, preferably at least an hour, five hours, 12 hours, a day, a week, a month, more preferably several months (e. g. up to three months), prior or subsequent to administration of composition of the present invention.

In some embodiments, the single variable heavy chain domain antibody that binds to MSLN, a binding molecule comprising a single variable heavy chain domain antibody that binds to MSLN or pharmaceutical composition described herein may be administered with two or more therapeutic agents. In some embodiments, the binding agents of the invention may be administered with two or more therapeutic agents.

- 10 The single variable heavy chain domain antibody that binds to MSLN, a binding molecule comprising a single variable heavy chain domain antibody that binds to MSLN or a pharmaceutical composition as described herein may be administered at the same time or at a different time as the other therapy or therapeutic compound or therapy, e.g., simultaneously, separately or sequentially.

Methods of treatment

MSLN is expressed on the surface of tumour cells and high expression levels of soluble MSLN have been correlated with poor prognosis in several cancers. Anti- MSLN antibodies have been investigated as anti-cancer therapeutics. These anti-MSLN antibodies either induce direct cell killing through their ADCC activity or are used in the form of ADCs.

- 20 The molecules and cells described herein are therefore expected to find application in the treatment of disease, in particular cancer.

In one embodiment, the disease is a disease associated with expression of mesothelin. The molecules of the invention may preferentially bind to MSLN present on the surface of a cancer cell as compared to soluble MSLN. The cancer to be treated using an antibody molecule of the invention therefore preferably expresses, or has been determined to express, MSLN. More preferably, cells of the cancer to be treated comprise, or have been determined to comprise, MSLN at their cell surface, i.e. to comprise cell-surface bound MSLN. Methods for determining the presence of an antigen on a cell surface are known in the art and include, for example, flow cytometry.

- 30 In one embodiment, the disease is cancer and the invention thus relates to methods for the prevention and/or treatment of cancer, comprising administering to a subject a cell or cell population comprising a single domain antibody as described herein, said method comprising

administering, to a subject in need thereof, a pharmaceutically active amount of a cell and/or of a pharmaceutical composition of the invention.

The invention also relates to a single domain antibody as described herein for use in therapy. The invention also relates to a single domain antibody as described herein for use in the treatment of cancer. The invention also relates to the use of a single domain antibody as described herein in the manufacture of a medicament for the treatment of cancer.

10 The term "cancer" refers to a disease characterized by the uncontrolled growth of aberrant cells. Cancer includes all types of cancerous growths or oncogenic processes, metastatic tissues or malignantly transformed cells, tissues or organs irrespective of the histopathologic type or stage of invasiveness. Cancer cells can spread locally or through the bloodstream and lymphatic system to other parts of the body. The cancer may be a primary or a secondary cancer. Thus, an antibody molecule as described herein may be for use in a method of treating cancer in an individual, wherein the cancer is a primary tumour and/or a tumour metastasis.

The cancer to be treated using an antibody molecule of the invention may be a solid cancer. Examples of various cancers are described further herein and include, but are not limited to, mesothelioma, breast cancer, prostate cancer, ovarian cancer, cervical cancer, skin cancer, pancreatic cancer, colorectal cancer, renal cancer, liver cancer, brain cancer, lymphoma, leukemia, lung cancer and the like.

20 The phrase "disease associated with expression of mesothelin" includes, but is not limited to, a disease associated with expression of mesothelin or condition associated with cells which express mesothelin including, e.g., proliferative diseases such as a cancer or malignancy or a precancerous condition such as a mesothelial hyperplasia; or a noncancer related indication associated with cells which express mesothelin. Examples of various cancers that express mesothelin include but are not limited to, mesothelioma, lung cancer, ovarian cancer, pancreatic cancer, and the like.

In one embodiment, the cancer is selected from a haematological cancer or malignancy or a solid tumor. Hematologic cancers are cancers of the blood or bone marrow. Solid tumors are abnormal masses of tissue that usually do not contain cysts or liquid areas.

In one embodiment, the cancer is metastatic.

30 Cancers that may be treated by methods, uses and compositions described herein include, but are not limited to, cancer cells from the bladder, blood, bone, bone marrow, brain, breast, colon, esophagus, gastrointestinal, gum, head, kidney, liver, lung, nasopharynx, neck, ovary, prostate, skin, stomach, testis, tongue, or uterus. In addition, the cancer may specifically be of the following histological type, though it is not limited to these: neoplasm, malignant;

carcinoma; carcinoma, undifferentiated; giant and spindle cell carcinoma; small cell carcinoma; papillary carcinoma; squamous cell carcinoma; lymphoepithelial carcinoma; basal cell carcinoma; pilomatrix carcinoma; transitional cell carcinoma; papillary transitional cell carcinoma; adenocarcinoma; gastrinoma, malignant; cholangiocarcinoma; hepatocellular carcinoma; combined hepatocellular carcinoma and cholangiocarcinoma; trabecular adenocarcinoma; adenoid cystic carcinoma; adenocarcinoma in adenomatous polyp; adenocarcinoma, familial polyposis coli; solid carcinoma; carcinoid tumor, malignant; bronchiolo-alveolar adenocarcinoma; papillary adenocarcinoma; chromophobe carcinoma; acidophil carcinoma; oxyphilic adenocarcinoma; basophil carcinoma; clear cell

10 adenocarcinoma; granular cell carcinoma; follicular adenocarcinoma; papillary and follicular adenocarcinoma; nonencapsulating sclerosing carcinoma; adrenal cortical carcinoma; endometrioid carcinoma; skin appendage carcinoma; apocrine adenocarcinoma; sebaceous adenocarcinoma; ceruminous adenocarcinoma; mucoepidermoid carcinoma; cystadenocarcinoma; papillary cystadenocarcinoma; papillary serous cystadenocarcinoma; mucinous cystadenocarcinoma; mucinous adenocarcinoma; signet ring cell carcinoma; infiltrating duct carcinoma; medullary carcinoma; lobular carcinoma; inflammatory carcinoma; paget's disease, mammary; acinar cell carcinoma; adenosquamous carcinoma; adenocarcinoma w/squamous metaplasia; thymoma, malignant; ovarian stromal tumor, malignant; thecoma, malignant; granulosa cell tumor, malignant; and robblastoma, malignant;

20 sertoli cell carcinoma; leydig cell tumor, malignant; lipid cell tumor, malignant; paraganglioma, malignant; extra-mammary paraganglioma, malignant; pheochromocytoma; glomangiosarcoma; malignant melanoma; amelanotic melanoma; superficial spreading melanoma; malig melanoma in giant pigmented nevus; epithelioid cell melanoma; blue nevus, malignant; sarcoma; fibrosarcoma; fibrous histiocytoma, malignant; myxosarcoma; liposarcoma; leiomyosarcoma; rhabdomyosarcoma; embryonal rhabdomyosarcoma; alveolar rhabdomyosarcoma; stromal sarcoma; mixed tumor, malignant; mullerian mixed tumor; nephroblastoma; hepatoblastoma; carcinosarcoma; mesenchymoma, malignant; brenner tumor, malignant; phyllodes tumor, malignant; synovial sarcoma; mesothelioma, malignant; dysgerminoma; embryonal carcinoma; teratoma, malignant; struma ovarii, malignant;

30 choriocarcinoma; mesonephroma, malignant; hemangiosarcoma; hemangioendothelioma, malignant; kaposi's sarcoma; hemangiopericytoma, malignant; lymphangiosarcoma; osteosarcoma; juxtacortical osteosarcoma; chondrosarcoma; chondroblastoma, malignant; mesenchymal chondrosarcoma; giant cell tumor of bone; ewing's sarcoma; odontogenic tumor, malignant; ameloblastic odontosarcoma; ameloblastoma, malignant; ameloblastic fibrosarcoma; pinealoma, malignant; chordoma; glioma, malignant; ependymoma; astrocytoma; protoplasmic astrocytoma; fibrillary astrocytoma; astroblastoma; glioblastoma; oligodendroglioma; oligodendroblastoma; primitive neuroectodermal; cerebellar sarcoma; ganglioneuroblastoma; neuroblastoma; retinoblastoma; olfactory neurogenic tumor;

meningioma, malignant; neurofibrosarcoma; neurilemmoma, malignant; granular cell tumor, malignant; malignant lymphoma; Hodgkin's disease; Hodgkin's lymphoma; paragranuloma; malignant lymphoma, small lymphocytic; malignant lymphoma, large cell, diffuse; malignant lymphoma, follicular; mycosis fungoides; other specified non-Hodgkin's lymphomas; malignant histiocytosis; multiple myeloma; mast cell sarcoma; immunoproliferative small intestinal disease; leukemia; lymphoid leukemia; plasma cell leukemia; erythroleukemia; lymphosarcoma cell leukemia; myeloid leukemia; basophilic leukemia; eosinophilic leukemia; monocytic leukemia; mast cell leukemia; megakaryoblastic leukemia; myeloid sarcoma; and hairy cell leukemia.

- 10 In one aspect, where the binding agent binds MSLN and PSMA, the binding agent is used to target PSMA and treat prostate cancer.

In therapies of prostatic disorders, e.g., prostate cancer, the therapy can be used in combination with existing therapies. In one embodiment, the binding agent is used in combination with an existing therapy or therapeutic agent, for example an anti-cancer therapy. Thus, in another aspect, the invention also relates to a combination therapy comprising administration of the binding agent or a pharmaceutical composition of the invention and an anti-cancer therapy. The anti-cancer therapy may include a therapeutic agent or radiation therapy and includes gene therapy, viral therapy, RNA therapy bone marrow transplantation, nanotherapy, targeted anti-cancer therapies or oncolytic drugs. Examples of other therapeutic agents include other checkpoint inhibitors, antineoplastic agents, immunogenic agents, attenuated cancerous cells, tumor antigens, antigen presenting cells such as dendritic cells pulsed with tumor-derived antigen or nucleic acids, immune stimulating cytokines (e.g., IL-2, IFN α 2, GM-CSF), targeted small molecules and biological molecules (such as components of signal transduction pathways, e.g. modulators of tyrosine kinases and inhibitors of receptor tyrosine kinases, and agents that bind to tumor- specific antigens, including EGFR antagonists), an anti-inflammatory agent, a cytotoxic agent, a radiotoxic agent, or an immunosuppressive agent and cells transfected with a gene encoding an immune stimulating cytokine (e.g., GM-CSF), chemotherapy. In one embodiment, the the binding agent or pharmaceutical composition of the invention is used in combination with surgery. The binding agent or pharmaceutical composition of the invention may be administered at the same time or at a different time as the other therapy, e.g., simultaneously, separately or sequentially.

In one embodiment, an immune checkpoint inhibitor is also administered with the cell or cell population or pharmaceutical composition. The immune checkpoint inhibitor may be an anti-PD1, anti PDL-1, anti PDL-2, anti CTL-4, anti-TIM-3 or anti LAG-3 antibody. In another embodiment, the immune checkpoint inhibitor is selected from nivolumab, pembrolizumab,

cemiplimab, avelumab, durvalumab, or atezolizumab, Spartalizumab, Camrelizumab, Sintilimab, Tislelizumab, Pidilizumab, Toripalimab, Ipilimumab or Tremelimumab. In another embodiment, the immune checkpoint inhibitor is an interfering nucleic acid molecule, a small molecule or a PROteolysis TARgeting Chimera (PROTAC).

The immune checkpoint inhibitor is administered before, after or at the same time as the cell or cell population.

Immunoconjugates and other agents

10

In another aspect, there is provided an immunoconjugate comprising a single variable heavy chain domain antibody that binds to MSLN or a binding molecule comprising a single variable heavy chain domain antibody that binds to MSLN described herein conjugated to at least one therapeutic and/or diagnostic agent.

The invention also relates to the use of a single variable heavy chain domain antibody that binds to MSLN or a binding molecule comprising a single variable heavy chain domain antibody that binds to MSLN described herein for use as a diagnostic agent. The invention also relates to the use of a single variable heavy chain domain antibody that binds to MSLN or a binding molecule comprising a single variable heavy chain domain antibody that binds to MSLN described herein conjugated to a label.

20

Kits and methods

In another aspect, the invention provides a kit for detecting cancer, treatment, prognosis or monitoring comprising a genetically modified cell or pharmaceutical composition of the invention. The kit may also comprise instructions for use. In one embodiment, the single domain antibody or pharmaceutical composition comprises a label and one or more compounds for detecting the label. The invention in another aspect provides a binding molecule of the invention packaged in lyophilized form, or packaged in an aqueous medium.

30

Exemplary non- therapeutic applications

In another aspect, a single variable heavy chain domain antibody that binds to MSLN described herein is used for non-therapeutic purposes, such as diagnostic tests and assays. A method for detecting the presence of human MSLN in a test sample comprises contacting said sample with a single domain antibody described herein and at least one detectable label and detecting binding of said single domain antibody to human MSLN.

Modifications of antibodies for diagnostic purposes are well known in the art. For example, antibodies may be modified with a ligand group such as biotin, or a detectable marker group such as a fluorescent group, a radioisotope, or an enzyme. Compounds of the invention can be used for diagnostic purposes and e.g. labelled using conventional techniques. Suitable detectable labels include but are not limited to fluorophores, chromophores, radioactive atoms, electron-dense reagents, enzymes, and ligands having specific binding partners.

10 Unless otherwise defined herein, scientific and technical terms used in connection with the present disclosure shall have the meanings that are commonly understood by those of ordinary skill in the art. While the foregoing disclosure provides a general description of the subject matter encompassed within the scope of the present invention, including methods, as well as the best mode thereof, of making and using this invention, the following examples are provided to further enable those skilled in the art to practice this invention and to provide a complete written description thereof. However, those skilled in the art will appreciate that the specifics of these examples should not be read as limiting on the invention, the scope of which should be apprehended from the claims and equivalents thereof appended to this disclosure. Various further aspects and embodiments of the present invention will be apparent to those skilled in the art in view of the present disclosure.

20 All documents and sequence database identifiers mentioned in this specification are incorporated herein by reference in their entirety.

"and/or" where used herein is to be taken as specific disclosure of each of the two specified features or components with or without the other. For example, "A and/or B" is to be taken as specific disclosure of each of (i) A, (ii) B and (iii) A and B, just as if each is set out individually herein. Unless context dictates otherwise, the descriptions and definitions of the features set out above are not limited to any particular aspect or embodiment of the invention and apply equally to all aspects and embodiments which are described.

The invention is further described in the non-limiting examples.

Examples

30 Materials and methods

Construction of Tg/TKO mice

Mice carrying a heavy-chain antibody transgenic locus in germline configuration within a background that is silenced for endogenous heavy and light chain antibody expression (triple knock-out, or TKO) were created as previously described (WO2004/076618 and WO2003/000737, Ren *et al.*, Genomics, 84, 686, 2004; Zou *et al.*, J. Immunol., 170, 1354,

2003, Teng et al³⁰ all incorporated herein by reference). Briefly, transgenic mice were derived following pronuclear microinjection of freshly fertilised oocytes with a yeast artificial chromosome (YAC) comprising multiple human V_H, D and J genes in combination with mouse immunoglobulin constant region genes lacking CH1 domains, mouse enhancer and regulatory regions. Yeast artificial chromosomes (YACs) are vectors that can be employed for the cloning of very large DNA inserts in yeast. As well as comprising all three cis-acting structural elements essential for behaving like natural yeast chromosomes (an autonomously replicating sequence (ARS), a centromere (CEN) and two telomeres (TEL)), their capacity to accept large DNA inserts enables them to reach the minimum size (150 kb) required for chromosome-like stability and for fidelity of transmission in yeast cells. The construction and use of YACs is well known in the art (e.g., Bruschi, C.V. and Gjuracic, K. Yeast Artificial Chromosomes, Encyclopedia of Life Sciences, 2002, Macmillan Publishers Ltd., Nature Publishing Group / www.els.net).

The YAC used comprised multiple human heavy chain V genes, human heavy chain D and J genes. It lacks the C_H1 exon.

The transgenic founder mice were back crossed with animals that lacked endogenous immunoglobulin expression to create the Tg/TKO lines used for immunisation with recombinant MSLN antigen.

Generation of V_H domains

EXAMPLE 2. Immunisation Protocol to generate MSLN binders

DNA immunisations were carried out at Aldevron www.aldevron.com

Full length mesothelin cDNA was cloned into a proprietary expression vector with an expression tag. Plasmid DNA was transiently transfected into cells and surface expression confirmed by flow cytometry, detecting the tag.

MSLN is synthesized as a 71-kD precursor protein, then cleaved by the endoprotease furin into:

- secreted N-terminal region (a.a. 1-295), called megakaryocyte potentiating factor (MPF)
- 41-kD mature MSLN (296-606) cell surface protein, which remains tethered to the membrane via a GPI link
- The full length protein is 622 amino acids, 607-622 are lost during processing of the GPI anchor

Eight Tg/TKO mice aged 8–12 weeks of age at the initiation of immunisation each received eight weekly doses of MSLN plasmid DNA delivered via Genegun. On day 59, sera from immunised animals were tested by flow cytometry using mammalian cells transiently

transfected with the target cDNA cloned into an Aldevron proprietary expression vector containing a N-terminal tag-sequence. As a negative control non-transfected cells were used that does not express the antigen of interest. Reactivity of the immune sera against cells transiently transfected with the test construct could be detected in the immunized animals when compared to negative control cells.

For immunisations, isolation and characterisation of PSMA binders see WO 2017/122017 incorporated herein by reference.

EXAMPLE 3. Serum ELISA

10 At the end of the immunisation, terminal bleeds were collected from all animals, processed to serum and assayed for the presence of heavy-chain antibody responses to the immunogen by ELISA.

Multiwell plates were coated with a His-tagged human or cyno mesothelin recombinant protein then washed with PBS. Non-specific protein interactions were blocked with 3% (w/v) skimmed milk powder (Marvel®) in PBS. Dilutions of serum in 3% Marvel™/PBS were incubated for one hour at room temperature then transferred to the blocked ELISA plate for at least one hour. Unbound protein was removed by repetitive washes with PBS/Tween20 followed by PBS. Biotin-conjugated, goat anti-mouse IgG, Fcγ subclass 1 specific antibody prepared in PBS/3% Marvel was added to each well and incubated at room temperature for one hour, then washed as above. Neutravidin-HRP solution in 3% Marvel/PBS was added to the ELISA plates
20 for 30 minutes, then washed as above and developed using TMB substrate. The reaction was stopped after 10 minutes by the addition of 0.5M sulphuric acid solution. Absorbances were determined by reading at an optical density of 450nm.

ELISA results showed that all animals had generated an immune response to human mesothelin. In addition, all animals showed cross reactivity to cyno mesothelin.

EXAMPLE 4. Generation of Libraries from Immunised Mice

a) processing tissues, RNA extraction and cDNA synthesis

Inguinal and axillary lymph nodes and spleens were collected from each immunised animal into RNeasy®. At the point of processing, tissues were removed from the RNeasy and placed in a microtube with a stainless steel bead and RNeasy Lysis Reagent. Tissues were
30 lysed and homogenised by physical disruption via shaking for 3 mins at 1600 rpm in a MPBio FastPrep96 homogeniser. The lysate was cleared by centrifugation, chloroform added, mixed by shaking, then separated into phases by chilled centrifugation at 4700rpm for 20 minutes. The aqueous phase was collected in a semi-automated way using the QIAcube robot. RNA

was prepared using RNeasy 96 QIAcube kit and QIAcube HT plastic ware, based on the manufacturer's protocol with minor modifications.

RNA quality was assessed using the QIAxcel electrophoresis automated DNA and RNA analysis system, running the RNA alongside the QX RNA 15nt alignment marker. RNA extracted from all spleen and lymph node tissues was found to be of high quality.

cDNA was synthesised using Superscript III RT-PCR high-fidelity kit (Invitrogen), following the manufacturer's guidelines. Five RT-PCR reactions were performed on each RNA sample, using a common reverse primer in combination with forward primers designed to specifically amplify VH from the specific frameworks present in the Crescendo mouse. cDNA products of the correct size were confirmed by analysis on the QIAxcel. cDNAs derived from lymph nodes from a single mouse were pooled together as were cDNAs from the spleen, then each pool of cDNAs was purified with the GeneJet PCR purification kit.

b) Cloning into phagemid vector

cDNA pools from spleen and lymph nodes were cloned into clean, linear phagemid vector pUCG3 using a PCR-based method.

800ng linearised pUCG3 was mixed with 200ng VH cDNA in a final volume of 50µl, including 1.5µl DMSO and 25µl Phusion GC 2x mix.

PCR was performed as follows:

	98°C 30sec		
20	98°C 10sec	}	2 cycles
	58°C 20sec		
	72°C 2min		
	72°C 5min		
	Hold at 10°C		

PCR products (cloned VH) were analysed on a 1% (w/v) agarose gel.

c) Creation of phagemid libraries in *E. coli*

Cloned V_H/phagemid PCR products were pooled, combining material from spleen and lymph nodes, in order to create one phage library per immunised animal. Material was purified using Fermentas PCR purification kit.

Phagemid DNA as above was mixed with 160µl of TG1 *E. coli*, split between two BioRad 0.2cm cuvettes and transformed by electroporation at 2500V, 25mF, 200W. Electroporated cells

were recovered in 10ml media for 1 hour at 37°C with shaking. A 10-fold dilution series of an aliquot of the transformed culture was plated onto Ampicillin agar plates and used to estimate library size. The remainder was plated on large format Bioassay dishes containing 2xTY agar supplemented with 2% (w/v) glucose and 100µg/ml ampicillin. All agar plates were incubated overnight at 30°C. Colonies were scraped into 10 ml of 2xTY broth; aliquots were stored at minus 80°C in cryovials after addition of 50% (v/v) glycerol solution.

All library sizes were in excess of 10^7 (range 2.6×10^7 to 4.9×10^7).

EXAMPLE 6. Selection strategies for isolation of MSLN binders

Preparation of library phage stocks and phage display selections were performed essentially according to published methods (Antibody Engineering, edited by Benny Lo, chapter 8, p161-176, 2004).

Round 1: panning selections were carried out on 10mg/mL human mesothelin recombinant protein (296-580) with C terminal His tag.

Round 2: panning selections were carried out on 10mg/mL cyno mesothelin recombinant protein (296-580) with C terminal His tag.

A number of colonies from R2 outputs were assayed for binding to human and cyno mesothelin by phage ELISA (40-48 colonies per output). Binding to human 5T4-his (negative control) was also assayed.

E. coli colonies were picked into liquid culture and grown overnight. Phage rescue was carried out by adding M13 K07, releasing phage displaying VH into the supernatant. ELISA plates were coated with the same antigens used for phage selection, then blocked as described in Example 3. Pre-incubated phage were added to the plate and allowed to bind antigen before multiple washes. Phage were bound by a secondary antibody, anti-M13-HRP, then detected with TMB as previously described and absorbance at 450nm was measured.

All outputs contained clones binding to both human and cyno mesothelin. There was no binding to human 5T4-His, as expected.

EXAMPLE 7. Sequencing MSLN binders from Round 2 outputs and identification of clones

All clones found to cross-bind human and cyno mesothelin were sequenced. 76 clones returned full length high quality sequence. One of these clones was MSLN V_H1.2, from which the final clone MSLN V_H1.1 was derived.

Amino acid sequence of MSLN V_H1.2 SEQ ID NO. 504

EITLKESGPTLVKPTQTLTLTCTFSGFSLSTSGLGVGWIRQPPGKALEWLALIYWND DKRYR
PSLKNRLTIKDTSKNQVVLMTNMDPVD TARYYCAHYNTSSETAFDIRGQGTMVTVSS

EXAMPLE 8. Design, cloning and small scale purification of clones that bind MSLN

Clone V_H1.2 was engineered in silico to remove a potential N-glycosylation site by substituting a serine residue (S) for the asparagine residue (N) at position 96 using Kabat numbering to create clone V_H1.1. The amino acid sequence of V_H1.1 is shown below, with the engineered amino acid underlined.

Amino acid sequence of V_H1.1 SEQ ID NO. 3

10 EITLKESGPTLVKPTQTLTLTCTFSGFSLSTSGLGVGWIRQPPGKALEWLALIYWND DKRYR
PSLKNRLTIKDTSKNQVVLMTNMDPVD TARYYCAHYSTSSETAFDIRGQGTMVTVSS

V_H1.1 was synthesised by TWIST Bioscience www.twistbioscience.com, custom cloned into *E. coli* expression vector pJex401 and sequence verified. For ease of purification, a C terminal 6-His tag was included.

Plasmid DNA was transformed into *E. coli* TG1 strain, cultured in a 96 deep well plate at 37°C in TB medium with 50mg/mL kanamycin, with shaking, to an OD₆₀₀ of 0.5-1. Protein expression was induced with IPTG at reduced temperature for approximately 16 hours.

20 Bacterial cells were pelleted by centrifugation at 4500rpm for 40 minutes and the supernatant recovered and filtered using a 0.45um membrane. Protein was bound to nickel resin in a sodium phosphate/sodium chloride buffer with 20mM imidazole, removing the flow through by applying a vacuum pressures of -5 kPa via a multi-well plate vacuum manifold, then washed. Protein was eluted by increasing addition of buffer with 200mM imidazole and centrifugation at 500g for 2 minutes, then buffer exchanged into PBS to a final concentration of 10mM.

EXAMPLE 9. Stability and binding kinetics of MLN094-B09

Overnight stability

30 100µL aliquots of V_H1.1 protein were dispensed into two identical 96 well plates and sealed with foil seals. One plate was incubated at 4°C and the other at 40°C for approximately 17 hours. Both samples were analysed by SE-UPLC with positive and negative controls. The SEC column (Waters ACQUITY) was run isocratically in SEC Buffer (5% 1-propanol, 200mM NaCl, 100mM sodium phosphate, pH 7.4) at 0.4 mL/min for 6 minutes per sample. Data was collected using a PDA detector at 280 nm. V_H1.1 protein was found to be 96.7% monomer after incubation at 4°C and 98.1% monomer after incubation at 40°C.

Binding kinetics to human mesothelin recombinant protein

Human mesothelin recombinant protein with a C-terminal His tag at 2 mg/mL was immobilised by amine coupling onto a CM5 sensor chip for 90 s at 25°C. The chip was quality controlled using dilutions of a MSLN binding V_H1.1 from 0.1 – 1000 nM. Protein samples prepared as described in Example 8 were normalised to 10mM, then analysed at 3 nM, 15 nM, 75 nM and 375 nM. Association and dissociation times were 180 s and 400 s respectively and the chip was regenerated for 20 s.

V_H1.1 was run using single cycle kinetics and a 1:1 binding model measured the binding kinetics to human mesothelin as:

$$k_a = 1.45 \times 10^6$$

$$10 \quad k_d = 4.49 \times 10^{-2}$$

$$K_D = 3.09 \times 10^{-8} \text{ (30.9 nM)}$$

Binding kinetics to cyno mesothelin recombinant protein

Cyno mesothelin recombinant protein with a C-terminal His tag at 2 mg/mL was immobilised by amine coupling onto a CM5 sensor chip for 150 s at 25°C. The chip was quality controlled essentially as described above. Protein samples prepared as described in Example 8 were normalised to 10mM, then analysed at 33.3 nM, 100 nM, 300 nM and 900 nM. Association and dissociation times were 180 s and 600 s respectively and the chip was regenerated for 2 x 30 s.

20 V_H1.1 was run using single cycle kinetics and a 1:1 binding model measured the binding kinetics to human cyno mesothelin as:

$$k_a = 1.07 \times 10^6$$

$$k_d = 4.68 \times 10^{-2}$$

$$K_D = 4.38 \times 10^{-8} \text{ (43.8 nM)}$$

The amino acid sequence of cyno MSLN is shown below (SEQ ID NO 507).

30 MALPMARPLSGSCGTPALGSLLFLLFSLGWVQPSRVLAGETRQEAAPLDGILTNPDIASLS
PRQLLGFTCVEVSGLSTELVQELAVAGQKNVKLSAEQLRCLAHRLSEPPEDLDALPLDLLL
FLNPDAFSGPQACTHFFSRVAKANVDLLPRGAPERQRLLPAALTCWGVRGSLSEADVRL
GGLACDLPGRFVAESAEEVLPRLVRCLGPLDQDQQAARAALQRRGGPPYGPSTWSISTL
DDLQSLLPVLGQPVIHSIPQGILAAWRQRSSRDPSWQQPEQTVLRPRFRDVERTTCPEEK
EVHEIDESLIFYKKRELEACVDAALLAAQMDRVD AIPFTYEQLDVLKHKLDELYPQGYPESVI
RHLGHLFLKMSPEDIRKWNVTSLET LKALLK VSKGHEMSAQVATLIDRVVVG RGQLDKDTA
DTLTAFCPGCLCSLSPERLSSVPPSIIGAVRPQDLDT CGPRQLDVLYPKARLAFQNMSGSEY
FVKIRPFLGGAPTE DLKALSQQNVSM DLATFMKLRREAVLPLSVAEVQKLLGPHVEGLKVEE

QHSPVRDWILKQRQDDLDLGLGLQGGIPNGYLILDLVREALSGTPCLLGPGPVLTVLALLL
ASTLA

Human and cyno MSLN have 88.42% identity across the full length sequence (622 amino acids).

EXAMPLE 10. FMAT Direct cell Binding Assay to test MSLN binders

V_H1.1 proteins purified as described in Example 8 were assayed using Fluorescence Microvolume Assay Technology (FMAT), a fluorescence-based platform that detects fluorescence localized to beads or cells settled at the bottom of microwells (Dietz *et al.*, *Cytometry* 23:177-186 (1996), Miraglia *et al.*, *J. Biomol. Screening* 4:193-204 (1999). CHO TREX cell lines were generated in-house using full-length human MSLN using standard procedures. Parent CHO cells (no modification; not expressing human MSLN) were used as a negative control.

Cells were resuspended at 1×10^5 cells/ml in FMAT assay buffer (pH 7.4) with the addition of 120nM DRAQ5. V_H proteins were diluted as a 10-point series ranging from 3.33×10^{-8} M to 1.69×10^{-12} M. 10 μ l per well was transferred into 384 well black clear-bottomed assay plates with 10 μ l of 6nM mouse Anti-His and 10 μ l of 12nM Goat Anti-Mouse Alexa Fluor-488. DRAQ5 stained cells (20 μ l per well) were added and the assay plates incubated in the dark for 2 hours at room temperature. Plates were read in the FL2 (502nm-537nm) and FL5 (677-800nm) channels on the TTP Mirrorball plate reader following excitation at 488nm and 640nm. Data was gated on FL5 perimeter and peak intensity and the FL2 median mean fluorescence intensity of the gated data used for determination of V_H binding.

The EC₅₀ of V_H1.1 was found to be 1.2 nM

EXAMPLE 11. Scale up purification of V_H1.1

MSLN V_H1.1 was cloned into pJex401 E. coli expression plasmid with a C terminal 6xHis tag spaced with a cleavable TEV protease recognition site, then sequence verified. Plasmid DNA was transformed into TG1 E. coli cells and grown in 900ml TB medium with 50mg/mL kanamycin, shaking, to an OD₆₀₀ of 0.5-1. Protein expression was induced with IPTG at reduced temperature for approximately 16 hours.

Ni Sepharose Excel affinity resin was equilibrated from storage buffer into PBS in a 200 mL Econo-Column, and the volume adjusted to a 50% slurry. An appropriate volume of slurry was added to the culture supernatants which were left to mix on a rolling bed (33 rpm) for ≥ 1 hour. The supernatant/resin mixtures were poured into clean 25 mL Econo-Columns fitted with 250 mL funnels to collect the resin and bound protein. Collected resin was washed with 20 resin volumes of 20 mM sodium phosphate, 500mM NaCl, 20 mM imidazole, pH 7.4 before the

bound protein was eluted in fractions of 20 mM sodium phosphate, 500mM NaCl, 500 mM imidazole, pH 7.4. The absorbance at 280 nm of each elution fraction was measured using a nanodrop.

Preparative SEC was performed using a HiLoad 26/600 Superdex 75 pg column running isocratically in PBS pH 7.4 on an Akta system. The elution samples from Ni-affinity chromatography were loaded via a sample pump with a maximum injection volume of 13 mL per run and eluted with 1.2 CV of PBS (pH 7.4) with a flow rate of 2.6 mL/min. The peak collection threshold was set at 20 mAU and 2 mL fractions were collected using a fraction collector. A280 was measured throughout the run using a UV detector. Fractions containing pure samples of protein were pooled and concentrated to 2mg/mL in a Amicon Ultra-15, 3,000 MWCO RC centrifugal filter unit, centrifuged at max. 4,000 xg.

The expression titre was 32.0 mg/L and 24.6mg of high purity protein was recovered.

EXAMPLE 12. Kinetic and affinity assessment by BiaCore of MSLN binders

Human mesothelin recombinant protein with a C-terminal His tag at 2 mg/mL was immobilised by amine coupling as described in Example 9. Association and dissociation times were 180 s and 400 s respectively and the chip was regenerated for 20 s.

V_H1.1 prepared as described in Example 11 was run using multi cycle kinetics using a 6-point dilution series ranging from 1000nM to 4.12 nM. Each sample was run twice.

1:1 binding model measured the binding kinetics to human mesothelin as:

$$k_a = 1.15\text{--}1.6 \times 10^6$$

$$k_d = 3.93\text{--}4.49 \times 10^{-2}$$

$$K_D = 2.8\text{--}3.4 \times 10^{-8} \text{ (28-34 nM)}$$

Name	K _a (1Ms)	k _d	K _D (M)	K _D (nM)
1.1	1,72E+06	3,45E-02	2.01E-08	20.1
1.2	1.45E_06	4,49E-02	3.09E-02	30.9

Example 16 CAR

CAR construction

The following antigen-binding moieties were used: scFv derived from the J591 Ab specific for PSMA; human V_H domain specific for PSMA (PSMA-V_H); scFv derived from a MSLN-specific Ab Amatuximab; human V_H domain specific for MSLN (MSLN-V_H). All ligands were assembled with the CD8α hinge and transmembrane domain, the CD28 costimulatory domain and CD3ζ

intracellular signaling domain and cloned into the SFG retroviral vector.²⁴ A FLAG-tag was incorporated after the antigen ligand to detect the expression of CARs by an anti-FLAG Ab. Dual specific (PSMA and MSLN) CARs were also generated by linking the two V_H domains. The linkers used are described in more detail below. The corresponding CARs were called J591, PSMA-V_H, MSLN scFv, MSLN-V_H and PSMA-V_H/MSLN-V_H. Retroviral supernatants were produced by transfection of 293 T cells with the retroviral vectors, the RD114 envelope from RDF plasmid and the MoMLV gag-pol from PegPam3-e plasmid. Supernatants were collected 48 hours and 72 hours after the transfection and filtered with 0.45 µm filter.²⁴

The polypeptide sequences of the V_H domains used were as follows:

10 MSLN-V_H SEQ ID NO. 3

PSMA-V_H SEQ ID NO. 8

PSMA-V_H- (G₄S)₃ -MSLN-V_H SEQ ID NO. 505

EVQLVESGGGVVQPGRSLRLSCAASGFSFSGYGMHWVRQAPGKGLEWYAYISYDGSNKY
YADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAKDPAWGLRLGESSYDFDIWGQ
GTMVTVSSGGGGSGGGGSGGGGSEITLKESGPTLVKPTQTLTLTCTFSGFSLSTSGLGVG
WIRQPPGKALEWLALIYWDDKRYRPSLKNRLTIAKDTSKNQVVLMTNMDPVDTARYYCA
HYSTSSETAFDIRGQGTMVTVSS

PSMA-V_H- (G₄S)₆ -MSLN-V_H SEQ ID NO. 506

20 EVQLVESGGGVVQPGRSLRLSCAASGFSFSGYGMHWVRQAPGKGLEWYAYISYDGSNKY
YADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAKDPAWGLRLGESSYDFDIWGQ
GTMVTVSSGGGGSGGGGSGGGGSGGGGSGGGGSGGGGSEITLKESGPTLVKPTQTLTL
TCTFSGFSLSTSGLGVGWIRQPPGKALEWLALIYWDDKRYRPSLKNRLTIAKDTSKNQVVL
TMTNMDPVDTARYYCAHYSTSSETAFDIRGQGTMVTVSS

Cell lines

Tumor cell lines PC-3, C4-2 (prostate cancer) and Aspc-1 (pancreatic cancer) were purchased from ATCC (American Type Culture Collection). All tumor cell lines were cultured with RPMI-1640 (Gibco) supplemented with 10% Fetal bovine serum (Sigma), 2 mM GlutaMax (Gibco) and penicillin (100 units/mL) and streptomycin (100 µg/mL; Gibco). All cells were cultured at 37°C with 5% CO₂. PC-3 cell line was transduced with retroviral vectors encoding PSMA or
30 MSLN to make PC-3-PSMA and PC-3-MSLN. PC-3-PSMA, PC-3-MSLN and Aspc-1 were transduced with retroviral vectors encoding Firefly-Luciferase-eGFP (FFluc-eGFP) gene.

CAR-T cell generation

Buffy coats from healthy donors (Gulf Coast Regional Blood Center) were processed with Lymphoprep density separation (Fresenius Kabi Norge) to isolate peripheral blood mononuclear cells, which were then activated on plates coated with 1 µg/mL CD3 (Miltenyi Biotec) and 1 µg/mL CD28 (BD Biosciences) monoclonal Abs (mAbs). Two days later, activated T cells were transduced with retroviral supernatants on 24-well plates coated with retronectin (Takara Bio). T cells were collected 3 days after transduction and expanded in 40% RPMI-1640(Gibco) and 40% Click's medium (Irvine Scientific), 10% HyClone FBS (GE healthcare), 2 mM GlutaMAX(Gibco), 100 unit/mL of Penicillin and 100 mg/mL of streptomycin (Gibco) with 10 ng/mL IL-7 (PeproTech) and 5 ng/mL IL-15 (PeproTech). T cells were collected
 10 for functional assays 12–14 days after activation.^{25 26}

Flow cytometry

mAbs for human CD3 (APC-H7; SK7; 560176), CD4 (BV711; SK3; 563028), CD8 (APC; SK1; 340584), CD45RA (PE; HI100; 555489), CD45RO (BV786; UCHL1; 564290), CD69 (FITC; L78; 347823), CCR7 (FITC; 150503; FAB197F-100), PD-1(PE-Cy7; EH12.1;561272), Lag3(PE;T47-530;565616), FLAG (APC; L5; 637308), Granzyme-B (PE;GB11;561142) from BD biosciences and BioLegend were used. Samples were acquired with BD FACSCanto II or BD LSRFortessa. A minimum of 10 000 events were acquired for each sample and were analyzed using FlowJo 10 (FlowJo).

Western blot

20 CAR-T cells were incubated with 2 µg anti-FLAG Ab in 100 µL PBS for 20 mins on ice and then with 2 µg goat antimouse secondary Ab for another 20 mins on ice. Cells were then incubated in the 37°C water bath for the selected time points and then lysed with 2 x Laemmli buffer for 10 mins. Cell lysates were then separated in 4% to 15% 10 well SDS-PAGE gels and transferred to polyvinylidene difluoride membranes at 75V for 120 mins (Bio-Rad). Blots were examined for human CD3ζ (Santa Cruz Biotechnology), p-Y142 CD3ζ (Abcam), pan-ERK (BD Biosciences), and pan-Akt, p-S473 Akt, and p-T202/Y204 MAPK (Cell Signaling Technology) with 1:1000 dilution in 5% TBS-Tween milk. Membranes were incubated with HRP-conjugated secondary goat anti-mouse or goat anti-rabbit IgG (Santa Cruz) at a dilution of 1:3000 and imaged with the ECL substrate kit (Thermofisher) on the ChemiDoc MP System (Bio-Rad)
 30 according to the manufacturer's instructions.²⁶

Proliferation assay

T cells were labeled with 1.5 mM carboxyfluorescein diacetate succinimidyl ester (CFSE; Invitrogen) and plated with tumor cells at an effector to target (E:T) ratio of 1:1. CFSE signal dilution from gated T cells on day 5 was measured using flow cytometry.²⁶

In vitro cytotoxicity assay

Tumor cells were seeded in 24-well plates at a concentration of 2.5×10^5 cells/well overnight. CAR-T cells were added to the plate at an E:T of 1:5 without exogenous cytokines. Cocultures were analyzed 5–7 days following coculture to measure residual tumor cells and T cells by flow cytometry. Dead cells were recognized by Zombie Aqua Dye (Biolegend) staining while CAR-T cells were identified by CD3 staining and tumor cells by GFP. ²⁶ CD69, PD-1 and Lag3 expression was measured by flow cytometry from day 0 to day 5 each day after coculture of CAR-T cells with tumor cells. For the granzyme-B staining, Golgi protein inhibitor (BD Biosciences) was added on day 1 of coculture for 6 hours. Cocultures were then first stained with Zombie Aqua Dye (Biolegend) and CD3 mAb, followed by fixation/permeabilization solution (BD Biosciences). Intracellular staining of granzyme-B was then conducted.

Cytokine analysis

CAR-T cells (1×10^5 cells) were cocultured with 2.5×10^5 tumor cells in 24-well plates without exogenous cytokines. Supernatant was collected after 24 hours, and cytokines (interferon- γ (IFN- γ) and IL-2) were measured by using ELISA kits (R&D, Research And Development system) in duplicates following manufacturer's instructions. ²⁶

Expression and purification of recombinant proteins

A panel of recombinant proteins was produced, comprising bispecific (2V_H) proteins that bind both PSMA and MSLN, monospecific V_H protein binding PSMA, monospecific V_H protein binding MSLN and a control scFv protein based on Amatuximab. Bispecific protein was made in two formats, one with a short flexible linker (G4S)₃, another one with a long flexible linker (G4S)₆. Bispecific proteins were expressed in mammalian cells and purified by protein A binding. Monospecific proteins were His tagged at the C terminus, expressed in *Escherichia coli* and purified by His trap and size exclusion chromatography.

Binding and kinetic analysis

Binding analyses were performed at 25°C using Biacore 8K system. The instrument was run on 1 x HBS-EP⁺ (BR100669) buffer and the data were analyzed using Biacore Insight Evaluation software. Recombinant human MSLN was diluted to 2 μ g/mL in 10 mM sodium acetate buffer pH4.0 and immobilized on a CM5 sensor chip (contact time 120 s) using amine-coupling kit with accordance to the manufacturer's instructions. Humabody V_H samples were tested for binding at 5 concentrations 3.7 nM, 11.1 nM, 33.3 nM, 100 nM and 300 nM using multicycle kinetics method. Each sample was injected for 100 s at the flow rate 35 μ L/min and dissociated for 100 s. The antigen surface was regenerated by 20 s injection of 10 mM glycine pH 2.0. Recombinant human PSMA antigen with a human Fc tag was captured on a Protein G sensor. Humabody V_H samples were tested in Single-cycle kinetics mode at increasing concentrations of 2.22 nM, 6.67 nM, 20 nM and 60 nM with 90 s association and 600 s

dissociation time at the flow rate of 30 μ L/min. Buffer injections were made to allow for double-reference subtraction. The sensor surface was regenerated with 10 mM glycine pH1.5 (GE Healthcare BR100354). To detect dual binding to MSLN and PSMA, human PSMA antigen surface was captured as above. Bispecific PSMA-MSLN Humabody constructs were captured on the PSMA surface by injecting 100 nM of each sample for 100 s at 35 μ L/min flow rate. The capture was immediately followed by an injection of 300 nM recombinant human MSLN with 100 s contact time and 100 s dissociation. A PSMA-specific Humabody construct without a MSLN-binding arm was used as a control.

Xenograft murine models

- 10 NSG (NOD scid gamma mouse) mice (6–8 weeks old) were injected intravenously through tail vein with either PC-3-PSMA-FFluc-eGFP, or PC-3-PSMA-FFluc-eGFP and PC-3-MSLN-FFluc-eGFP mixed at 1 to 1 ratio, or Aspc-1-FFluc-eGFP tumor cells of 1×10^6 cells per mice. Fourteen days later, CAR-T cells were injected intravenously through tail vein. For the high dose treatment, 4×10^6 CAR-T cells per mice were injected, while for the low dose treatment, 1×10^6 CAR-T cells per mice were injected. In the rechallenge experiments, mice were infused 1×10^6 tumor cells per mice on clearance of the previous tumor. Tumor growth was monitored by bioluminescence using IVIS (In Vivo Imaging Systems)-Kinetics Optical in vivo imaging system (PerkinElmer) (PSMA- V_H and MSLN- V_H part) or AMI (AMI Medical Imaging) Optical in vivo imaging system (Spectral instruments imaging) (PSMA- V_H /MSLN- V_H part).

Statistics

All data was calculated and represented as mean with SD. One-way analysis of variance (ANOVA) or two-way ANOVA analyses were performed to compare multiple groups. Two-tailed t-test was used to compare two groups. P value of less than 0.05 was significant. All calculations and figures were achieved by GraphPad Prism V.7 (La Jolla, California, USA).

Results

Human V_H domain-based CAR targeting PSMA is expressed and signals in T cells

- 30 We constructed the PSMA-specific CARs using the scFv from the J591 mAb (J591) and the PSMA binding human V_H domain (PSMA- V_H) joined to the CD8 α stalk, CD28 costimulatory domain and CD3 ζ intracellular domain. A flag-based tag was incorporated into the cassettes to detect CAR expression by flow cytometry (figure 1A). Activated T cells were successfully transduced and expressed the CARs equally (figure 1B,C). The CD19-specific CAR (CD19) and non-transduced (NT) T cells were used as controls. On transduction, J591-T cells and PSMA- V_H -T cells showed similar expansion in vitro when exposed to IL-15 and IL-7 cytokines, which was similar to CD19-T cells and NT-T cells (figure 1D). Furthermore, no differences were observed in T cell composition as assessed by flow cytometry at day 12–14 of culture (figure

1E). We examined proximal signaling of CAR-T cells before and after CAR cross-linking mediated by an anti-Flag Ab. Phosphorylation of the CAR-associated CD3 ζ as well as phosphorylation of Akt and ERK were equal in J591-T cells and PSMA-V_H-T cells (figure 1F). Therefore, a V_H domain-based CAR is expressed and signals in T cells on cross-linking as observed for scFv-based CAR-T cells.

PSMA-specific V_H domain-based CAR-T cells are functional in vitro and in vivo

To compare the antitumor effect of PSMA-V_H-T cells and J591-T cells in vitro, we engineered the PC3 cells to express PSMA antigen (figure 2A), and we observed that J591-T cells and PSMA-V_H-T cells showed comparable granzyme-B expression when cultured with or without tumor cells (figure 2B, C and figure 8A). Similarly, both PSMA-V_H-T cells and J591-T cells showed equal upregulation and subsequent down regulation of CD69 as a marker of T cell activation (figure 2D). Similar upregulation on antigen stimulation and down regulation after antigen removal were observed for PD-1 and Lag-3 (figure 8B-E). We then cocultured PSMA-V_H-T cells and J591-T cells in vitro with tumor cells (either PSMA-positive or PSMA-negative), and measured the remaining tumor cells after 5 days of coculture. CD19-T cells did not eliminate tumor cells, while PSMA-V_H-T cells specifically eliminated PSMA-positive target cells (C4-2 and PC3-PSMA) to the same extent as conventional J591-T cells, and did not demonstrate off-target effect on PSMA-negative cells (PC3) (figure 2F,G). We also measured the secretion of IFN γ and IL-2 after 24 hours of coculture with tumor cells. When the PSMA-V_H-T and J591-T cells target PSMA-positive C4-2 and PC3-PSMA cells, both of them secreted high amount of IFN γ and IL-2 compared with control CD19-CAR-T cells (figure 2H,I). Furthermore, PSMA-V_H-T cells and J591-T cells proliferated similarly on encounter with tumor cells as shown by CFSE dilution (figure 2J). To investigate the antitumor effects of Humabody V_H CAR-T cells in vivo, NSG mice engrafted with PSMA-positive tumor cells labeled with Firefly luciferase were treated with a high doses (4×10^6 cells/mouse) of CAR-T cells (figure 3A). CAR-T cell treatment showed tumor control as measured by tumor bioluminescence without differences in mice treated with PSMA-V_H-T cells or J591-T cells (figure 3B,C). To further assess differences between PSMA-V_H-T cells and J591-T cells, we used low doses of T cells (1×10^6 cells/mouse) in tumor-bearing mice (figure 3D). We observed that PSMA-V_H-T cells still eliminated tumor cells in vivo as J591-T cells (figure 3E,F). In addition, we also observed similar V_H CAR-T cell persistence in the peripheral blood, spleen and bone marrow compared with traditional scFv-based CAR-T cells at day 58 at the time of euthanasia (figure 3G,H). Therefore, Humabody V_H CAR-T cells demonstrated comparable antitumor effects to scFv-based CAR-T cells in vitro and in vivo.

MSLN-specific V_H domain-based CAR-T cells demonstrate antitumor activity

To further assess the reproducibility of V_H domain-based CARs, we tested a MSLN-specific Humabody V_H. We constructed the conventional MSLN scFv CAR (MSLN-scFv) and V_H domain CAR (MSLN-V_H) using the same backbone developed for PSMA-specific CARs (figure 4A). MSLN-scFv and MSLN-V_H were equally expressed in T cells (figure 9A, B). In a similar fashion, we examined the antitumor activity of MSLN-scFv-T cells and MSLN-V_H-T cells in vitro through coculture experiments with tumor cells, cytokine release assay, proliferation assay. MSLN-scFv-T cells and MSLN-V_H-T cells selectively eliminated Aspc-1 tumor cells that express MSLN, while spared PC3 cells that do not express MSLN (figure 4B and figure 10). They released similar amount of IFN γ and IL-2 (figure 4C) and proliferated on encounter with tumor cells (figure 4D). In the xenotransplant model in NSG mice engrafted with Aspc-1 cells labeled with Firefly luciferase (figure 4E), MSLN-V_H-T cells showed even more profound antitumor effects as compared with mice treated with MSLN-scFv CAR-T cells (figure 4F,G), which translated in prolonged survival of the mice (figure 4H). However, we observed similar T cells expansion/persistence between MSLN-V_H and MSLN-ScFv (figure 4I). Thus V_H domain-based CARs can reproducibly redirect antitumor activity of engineered T cells.

In vitro analysis of monovalent and bivalent V_H domain recombinant proteins

To test whether the V_H domains are suitable to construct bispecific CARs, two V_H domains in tandem recombinant proteins linking PSMA-specific and MSLN-specific V_H were generated (figure 5A). To test whether the linkers had any effect on the target binding affinity, two different linkers were used: the (G4S)₃ linker ('short flexible linker') and a longer linker (G4S)₆ with 6 copies of the (G4S) repeat ('long flexible linker'). Monomer V_H proteins and a MSLN binding scFv were made as controls (figure 5A). Analysis of binding to PSMA recombinant protein by surface plasmon resonance (SPR) Biacore assay showed that the affinity of the PSMA-V_H remained the same when the PSMA-V_H was formatted with the MSLN-V_H domain using either flexible linkers (figure 5B). Similarly, analysis of binding to MSLN recombinant protein by SPR Biacore assay showed that the affinity of the MSLN-V_H domain was not altered when the PSMA-V_H was formatted with the MSLN-V_H using either flexible linkers (figure 5C). In summary, these data show that V_H modules in bispecific format are capable of binding their specific target with the same affinity as their monovalent counterparts.

Bispecific V_H domain-based CAR-T cells demonstrate dual specificity

We constructed a bispecific V_H domain CAR to facilitate CAR-T cells to specifically recognize two antigens simultaneously. We used the MSLN-V_H and PSMA-V_H domains fused with the short (G4S)₃ linker to generate the bispecific PSMA-V_H/MSLN-V_H CAR (figure 6A). The PSMA-V_H/MSLN-V_H CAR was expressed in T cells (figure 6B,C). PSMA-V_H-T cells, MSLN-V_H-T cells and PSMA-V_H/MSLN-V_H-T cells were cocultured with tumor the cell line Aspc-1, which express MSLN, and the PC3-PSMA cell line. We observed the PSMA-V_H/MSLN-V_H-T can eliminate

both tumor cell lines compared with single CAR-T cells, which only eliminate tumor cells expressing the targeted antigen (figure 11A, B). In addition, we also observed the expected cytokine release profile (figure 11C, D). Next, we confirmed that PSMA-V_H/MSLN-V_H-T cells displayed specific cytotoxicity toward the same cell line PC3 expressing either MSLN or PSMA similar to MSLN-V_H-T cells and PSMA-V_H-T cells without off-target effect (figure 6D,E). Importantly, when PC3-PSMA and PC3-MSLN were plated as 1:1 ratio mixture in coculture experiments, only PSMA-V_H/MSLN-V_H-T cells fully eliminated the tumor cells, although PSMA-V_H-T cells and MSLN-V_H-T cells showed some bystander killing effect as previously observed²⁷ (figure 6D,E). The in vitro antitumor effect was corroborated by release of IFN- γ and IL-2 (figure 6F,G). To evaluate if bispecific V_H domain CAR-T cells can eradicate tumors with mixed antigen expression in vivo, we established a metastatic xenograft mouse model by infusing PC3-PSMA cells and PC3-MSLN cells at 1:1 ratio into NSG mice by intravenous injection. Mice were then treated with CAR19-T, PSMA-V_H-T, MSLN-V_H-T and PSMA-V_H/MSLN-V_H-T cells (figure 7A). Dual targeting PSMA-V_H/MSLN-V_H-T cells controlled the tumor growth more effectively than either single targeting PSMA-V_H-T or MSLN-V_H-T cells (figure 7B,C). CAR-T cells were detectable in the peripheral blood of these mice up to 4 weeks after infusion (figure 7D). We also observed that T cells expressing bispecific CAR showed similar phenotypic profile as single CAR targeting T cells for exhaustion and memory markers (figure 12A-C). Analyses of antigen expression in tumor cells in vivo showed that tumor cells growing in mice receiving either PSMA-V_H-T or MSLN-V_H-T cells were predominantly MSLN and PSMA expressing cells, respectively (figure 7E). These results indicate that bispecific CARs generated by joining two human Ab V_H domains can prevent tumor escape in tumor with heterogeneous antigen expression.

Discussion

CARs approved by the Food and Drug Administration and those in clinical studies are mostly based on scFv-binding moieties. Here we demonstrated that monospecific human V_H domain-based CAR-T cells achieved comparable antitumor effects both in vitro and in vivo as scFv-based CAR-T cells. Furthermore, V_H domains combined in tandem to create bispecific molecules allowed the generation of effective CAR-T cells targeting two antigens.

Redirected T cell based on single-domain Abs have been recently proposed.^{17 28 29} However, most of them are obtained from llamas or camelid-derived libraries. Biological therapeutic molecules with non-human sequence can cause immune responses.^{18 28} Transgenic mouse technology has enabled the generation of biophysically robust fully human V_H domains known as Humabody V_H or Humabodies³⁰ which have the potential for use in CAR constructs while mitigating immunogenicity risk.

Despite the remarkable clinical activity of CAR-T cells in hematological malignancies, objective responses in patients with solid tumors are modest.^{10 26 31–33} Heterogeneity of antigen expression is one of the main reasons causing tumor escape in solid tumors after targeted therapies.^{10 19 20} Furthermore, murine-based scFv may cause immune responses especially in solid tumor patients who are usually less immunosuppressed compared with patients with liquid tumors. Targeting multiple TAAs and using human binding moiety in CAR molecules may improve the outcome of CAR-T cells in solid tumors.¹⁰ Here, we demonstrated that human V_H domains generated from a transgenic mouse might solve both issues of immunogenicity and tumor heterogeneity since bispecific CAR-T cells can be efficiently generated using two human V_H domains in tandem.

In addition to the issue of heterogeneity in antigen expression, the complex inhibitory pathways of the tumor microenvironment in solid tumors mean that additional genetic modification of T cells would likely be required to enhance T cell trafficking and functions.^{5 31 34–36} Generation of vector cassettes encoding multiple genes requires a significant optimization of the engineering strategies since the size of the entire cassette is limited. V_H domains are a good alternative to scFv since they are approximately half the size.

Here, we have used two target antigens, PSMA and MSLN, that are currently under evaluation to treat mesothelioma, lung cancer, breast cancer, pancreatic cancer and prostate cancer via scFv-based CAR-T cells.^{37–39} Our preclinical experiments validate the potential use of bispecific human V_H domains targeting both PSMA and MSLN in these difficult to treat malignancies. It remains to be validated if dual or multiple targeting with V_H domain-based CARs can be broadly applicable, and if targeting multiple antigens in solid tumors leads to increased potential for toxicity.

Additionally, we observed that V_H domain-based CAR-T cells have comparable cytotoxicity and proliferative capacity as traditional scFv-based CAR-T cells. MSLN-V_H-T cells showed even more profound antitumor effects as compared with mice treated with MSLN-scFv CAR-T cells. Interestingly, MSLN-V_H showed lower affinity than MSLN-scFv (28 nM compared with 79pM) recapitulating what has been observed for other scFvs that very high affinity is not necessarily optimal for CAR-based targeting for some targets.^{40–42} However, we cannot exclude that the observed superior antitumor activity of the MSLN-V_H-based CAR-T cells can be associated with the recognition of a different epitope rather than to different affinity. In summary, we have demonstrated that V_H domain CAR-T cells in monospecific format achieved comparable antitumor response compared with traditional scFv-based CAR-T cells both in vitro and in vivo. Furthermore, bispecific V_H domain CAR-T cells delivered potent anti-tumor effects demonstrating the potential to target solid tumors with heterogeneous antigen expression.

These proof-of-concept experiments lay the foundation for further development of human V_H domain-based CAR-T cells in clinical trials.

References

1. Eshhar Z , Waks T , Gross G , et al . Specific activation and targeting of cytotoxic lymphocytes through chimeric single chains consisting of antibody-binding domains and the gamma or zeta subunits of the immunoglobulin and T-cell receptors. *Proc Natl Acad Sci U S A* 1993;90:720–4.doi:10.1073/pnas.90.2.720
pmid:http://www.ncbi.nlm.nih.gov/pubmed/8421711
2. Finney HM , Lawson AD , Bebbington CR , et al . Chimeric receptors providing both primary and costimulatory signaling in T cells from a single gene product. *J Immunol* 1998;161:2791–7.pmid:http://www.ncbi.nlm.nih.gov/pubmed/9743337
3. Imai C , Mihara K , Andreansky M , et al . Chimeric receptors with 4-1BB signaling capacity provoke potent cytotoxicity against acute lymphoblastic leukemia. *Leukemia* 2004;18:676–84.doi:10.1038/sj.leu.2403302
pmid:http://www.ncbi.nlm.nih.gov/pubmed/14961035
4. Sadelain M , Brentjens R , Rivière I . The basic principles of chimeric antigen receptor design. *Cancer Discov* 2013;3:388–98.doi:10.1158/2159-8290.CD-12-0548
pmid:http://www.ncbi.nlm.nih.gov/pubmed/23550147
5. Dotti G , Gottschalk S , Savoldo B , et al . Design and development of therapies using chimeric antigen receptor-expressing T cells. *Immunol Rev* 2014;257:107–26.doi:10.1111/imr.12131 pmid:http://www.ncbi.nlm.nih.gov/pubmed/24329793
6. Xiong W , Chen Y , Kang X , et al . Immunological synapse predicts effectiveness of chimeric antigen receptor cells. *Mol Ther* 2018;26:963–75.doi:10.1016/j.ymthe.2018.01.020
pmid:http://www.ncbi.nlm.nih.gov/pubmed/29503199
7. Barber A , Zhang T , DeMars LR , et al . Chimeric NKG2D receptor-bearing T cells as immunotherapy for ovarian cancer. *Cancer Res* 2007;67:5003–8.doi:10.1158/0008-5472.CAN-06-4047 pmid:http://www.ncbi.nlm.nih.gov/pubmed/17510432
8. Shaffer DR , Savoldo B , Yi Z , et al . T cells redirected against CD70 for the immunotherapy of CD70-positive malignancies. *Blood* 2011;117:4304–14.doi:10.1182/blood-2010-04-278218
pmid:http://www.ncbi.nlm.nih.gov/pubmed/21304103

9. Brown CE , Alizadeh D , Starr R , et al . Regression of glioblastoma after chimeric antigen receptor T-cell therapy. *N Engl J Med* 2016;375:2561–9.doi:10.1056/NEJMoa1610497 pmid:http://www.ncbi.nlm.nih.gov/pubmed/28029927
10. Ahn S , Li J , Sun C , et al . Cancer immunotherapy with T cells carrying bispecific receptors that mimic antibodies. *Cancer Immunol Res* 2019;7:773–83.doi:10.1158/2326-6066.CIR-18-0636 pmid:http://www.ncbi.nlm.nih.gov/pubmed/30842091
11. Miller BR , Demarest SJ , Lugovskoy A , et al . Stability engineering of scFvs for the development of bispecific and multivalent antibodies. *Protein Eng Des Sel* 2010;23:549–57.doi:10.1093/protein/gzq028 pmid:http://www.ncbi.nlm.nih.gov/pubmed/20457695
12. Landoni E , Fucá G , Wang J , et al . Modifications to the framework regions eliminate chimeric antigen receptor tonic signaling. *Cancer Immunol Res* 2021. doi:doi:10.1158/2326-6066.CIR-20-0451. [Epub ahead of print: 05 Feb 2021].pmid:http://www.ncbi.nlm.nih.gov/pubmed/33547226
13. Holliger P , Hudson PJ . Engineered antibody fragments and the rise of single domains. *Nat Biotechnol* 2005;23:1126–36.doi:10.1038/nbt1142 pmid:http://www.ncbi.nlm.nih.gov/pubmed/16151406
14. Ahmad ZA , Yeap SK , Ali AM , et al . scFv antibody: principles and clinical application. *Clin Dev Immunol* 2012;2012:1–15.doi:10.1155/2012/980250 pmid:http://www.ncbi.nlm.nih.gov/pubmed/22474489
15. Hamers-Casterman C , Atarhouch T , Muyldermans S , et al . Naturally occurring antibodies devoid of light chains. *Nature* 1993;363:446–8.doi:10.1038/363446a0 pmid:http://www.ncbi.nlm.nih.gov/pubmed/8502296
16. Stanfield RL , Dooley H , Flajnik MF , et al . Crystal structure of a shark single-domain antibody V region in complex with lysozyme. *Science* 2004;305:1770–3.doi:10.1126/science.1101148 pmid:http://www.ncbi.nlm.nih.gov/pubmed/15319492
17. Jamnani FR , Rahbarizadeh F , Shokrgozar MA , et al . T cells expressing VHH-directed oligoclonal chimeric HER2 antigen receptors: towards tumor-directed oligoclonal T cell therapy. *Biochim Biophys Acta* 2014;1840:378–86.doi:10.1016/j.bbagen.2013.09.029 pmid:http://www.ncbi.nlm.nih.gov/pubmed/24076235
18. Lam N , Trinklein ND , Buelow B , et al . Anti-BCMA chimeric antigen receptors with fully human heavy-chain-only antigen recognition domains. *Nat Commun* 2020;11:283.

doi:10.1038/s41467-019-14119-9

pmid:http://www.ncbi.nlm.nih.gov/pubmed/31941907

19. O'Rourke DM , Nasrallah MP , Desai A , et al . A single dose of peripherally infused EGFRvIII-directed CAR T cells mediates antigen loss and induces adaptive resistance in patients with recurrent glioblastoma. *Sci Transl Med* 2017;9:eaaa0984. doi:10.1126/scitranslmed.aaa0984
pmid:http://www.ncbi.nlm.nih.gov/pubmed/28724573

20. Sampson JH , Gunn MD , Fecci PE , et al . Brain immunology and immunotherapy in brain tumours. *Nat Rev Cancer* 2020;20:12–25.doi:10.1038/s41568-019-0224-7
pmid:http://www.ncbi.nlm.nih.gov/pubmed/31806885

21. Chang SS . Overview of prostate-specific membrane antigen. *Rev Urol* 2004;6 Suppl 10:S13–18.pmid:http://www.ncbi.nlm.nih.gov/pubmed/16985927

22. Yu H , Pan J , Guo Z , et al . Cart cell therapy for prostate cancer: status and promise. *Onco Targets Ther* 2019;12:391–5.doi:10.2147/OTT.S185556
pmid:http://www.ncbi.nlm.nih.gov/pubmed/30655675

23. Morello A , Sadelain M , Adusumilli PS . Mesothelin-Targeted cars: driving T cells to solid tumors. *Cancer Discov* 2016;6:133–46.doi:10.1158/2159-8290.CD-15-0583
pmid:http://www.ncbi.nlm.nih.gov/pubmed/26503962

24. Hong LK , Chen Y , Smith CC , et al . CD30-Redirected Chimeric Antigen Receptor T Cells Target CD30+ and CD30- Embryonal Carcinoma via Antigen-Dependent and Fas/FasL Interactions. *Cancer Immunol Res* 2018;6:1274–87.doi:10.1158/2326-6066.CIR-18-0065 pmid:http://www.ncbi.nlm.nih.gov/pubmed/30087115

25. Wang Q , Chen Y , Park J , et al . Design and production of bispecific antibodies. *Antibodies* 2019;8:43.doi:10.3390/antib8030043

26. Du H , Hirabayashi K , Ahn S , et al . Antitumor responses in the absence of toxicity in solid tumors by targeting B7-H3 via chimeric antigen receptor T cells. *Cancer Cell* 2019;35:221–37.doi:10.1016/j.ccell.2019.01.002
pmid:http://www.ncbi.nlm.nih.gov/pubmed/30753824

27. Diaconu I , Ballard B , Zhang M , et al . Inducible caspase-9 selectively modulates the toxicities of CD19-specific chimeric antigen receptor-modified T cells. *Mol Ther* 2017;25:580–92.doi:10.1016/j.ymthe.2017.01.011
pmid:http://www.ncbi.nlm.nih.gov/pubmed/28187946

28. Rajabzadeh A , Rahbarizadeh F , Ahmadvand D , et al . A VHH-Based anti-MUC1 chimeric antigen receptor for specific retargeting of human primary T cells to MUC1-positive cancer cells. *Cell J* 2021;22:502–13.doi:10.22074/cellj.2021.6917
pmid:http://www.ncbi.nlm.nih.gov/pubmed/32347044
29. Sharifzadeh Z , Rahbarizadeh F , Shokrgozar MA , et al . Genetically engineered T cells bearing chimeric nanoconstructed receptors harboring TAG-72-specific camelid single domain antibodies as targeting agents. *Cancer Lett* 2013;334:237–44.doi:10.1016/j.canlet.2012.08.010
pmid:http://www.ncbi.nlm.nih.gov/pubmed/22902507
- 10 30. Teng Y , Young JL , Edwards B , et al . Diverse human VH antibody fragments with bio-therapeutic properties from the Crescendo Mouse. *N Biotechnol* 2020;55:65–76.doi:10.1016/j.nbt.2019.10.003
pmid:http://www.ncbi.nlm.nih.gov/pubmed/31600579
31. Fucà G , Reppel L , Landoni E , et al . Enhancing chimeric antigen receptor T-cell efficacy in solid tumors. *Clin Cancer Res* 2020;26:2444–51.doi:10.1158/1078-0432.CCR-19-1835 pmid:http://www.ncbi.nlm.nih.gov/pubmed/32015021
32. Shah NN , Fry TJ . Mechanisms of resistance to CAR T cell therapy. *Nat Rev Clin Oncol* 2019;3:372–85.doi:10.1038/s41571-019-0184-6
- 20 33. Ramos CA , Grover NS , Beaven AW , et al . Anti-Cd30 CAR-T cell therapy in relapsed and refractory Hodgkin lymphoma. *J Clin Oncol* 2020;38:3794–804.doi:10.1200/JCO.20.01342 pmid:http://www.ncbi.nlm.nih.gov/pubmed/32701411
34. Ma X , Shou P , Smith C , et al . Interleukin-23 engineering improves CAR T cell function in solid tumors. *Nat Biotechnol* 2020;38:448–59.doi:10.1038/s41587-019-0398-2
pmid:http://www.ncbi.nlm.nih.gov/pubmed/32015548
35. Bollard CM , Tripic T , Cruz CR , et al . Tumor-Specific T-cells engineered to overcome tumor immune evasion induce clinical responses in patients with relapsed Hodgkin lymphoma. *J Clin Oncol* 2018;36:1128–39.doi:10.1200/JCO.2017.74.3179
pmid:http://www.ncbi.nlm.nih.gov/pubmed/29315015
- 30 36. Heczey A , Louis CU , Savoldo B , et al . Car T cells administered in combination with Lymphodepletion and PD-1 inhibition to patients with neuroblastoma. *Mol Ther* 2017;25:2214–24.doi:10.1016/j.ymthe.2017.05.012
pmid:http://www.ncbi.nlm.nih.gov/pubmed/28602436
37. Wang D , Shao Y , Zhang X , et al . Il-23 and PSMA-targeted duo-CAR T cells in prostate cancer eradication in a preclinical model. *J Transl Med* 2020;18:23.

doi:10.1186/s12967-019-02206-w

pmid:http://www.ncbi.nlm.nih.gov/pubmed/31937346

38. Lv J , Zhao R , Wu D , et al . Mesothelin is a target of chimeric antigen receptor T cells for treating gastric cancer. *J Hematol Oncol* 2019;12:18.doi:10.1186/s13045-019-0704-y pmid:http://www.ncbi.nlm.nih.gov/pubmed/30777106
39. Zhang Z , Jiang D , Yang H , et al . Modified CAR T cells targeting membrane-proximal epitope of mesothelin enhances the antitumor function against large solid tumor. *Cell Death Dis* 2019;10:476.doi:10.1038/s41419-019-1711-1 pmid:http://www.ncbi.nlm.nih.gov/pubmed/31209210
- 10 40. Watanabe K , Terakura S , Uchiyama S , et al . Excessively high-affinity single-chain fragment variable region in a chimeric antigen receptor can counteract T-cell proliferation. *Blood* 2014;124:4799. doi:10.1182/blood.V124.21.4799.4799
41. Richman SA , Nunez-Cruz S , Moghimi B , et al . High-Affinity GD2-Specific CAR T cells induce fatal encephalitis in a preclinical neuroblastoma model. *Cancer Immunol Res* 2018;6:36–46.doi:10.1158/2326-6066.CIR-17-0211 pmid:http://www.ncbi.nlm.nih.gov/pubmed/29180536
42. Park S , Shevlin E , Vedvyas Y , et al . Micromolar affinity CAR T cells to ICAM-1 achieves rapid tumor elimination while avoiding systemic toxicity. *Sci Rep* 2017;7:14366. doi:10.1038/s41598-017-14749-3 pmid:http://www.ncbi.nlm.nih.gov/pubmed/29085043
- 20

CLAIMS

1. An isolated V_H single domain antibody that specifically binds to human mesothelin (MSLN), wherein the V_H single domain antibody comprising a CDR1 comprising SEQ NO. 4 or a sequence with 1, 2 or 3 amino acid modifications, CDR2 comprising SEQ NO. 5 or a sequence with 1, 2, 3 or 4 amino acid modifications and a CDR3 comprising SEQ ID NO. 6 or a sequence with 1, 2 or 3 amino acid modifications.
2. The isolated V_H single domain antibody according to claim 1 wherein the V_H single domain antibody comprises SEQ ID NO. 3 or a sequence having at least 75%, 80%, 90% or 95% sequence identity thereto.
3. The isolated V_H single domain antibody according to claim 1 or 2 wherein the antibody molecule is capable of activating an immune cell in the presence of MSLN.
4. The isolated V_H single domain antibody according to claim 3 wherein the immune cell is a T cell, B cell, natural killer (NK) cell, natural killer T (NKT) cell, or dendritic cell (DC).
5. The isolated V_H single domain antibody according to a preceding claim wherein the V_H single domain antibody binds to human MSLN with an affinity of 20 to 40 nM.
6. The isolated V_H single domain antibody according to a preceding claim wherein the V_H single domain antibody binds to cyno MSLN.
7. An isolated nucleic acid encoding a V_H single domain antibody according to any of claims 1 to 6.
8. A vector comprising a nucleic acid according to claim 7.
9. The vector according to claim 8 which is a retroviral vector, a DNA vector, a plasmid, a RNA vector, an adenoviral vector, an adenovirus associated vector, a lentiviral vector, or any combination thereof.
10. A host cell comprising a nucleic acid according to claim 7 or a vector according to claim 8 or 9.
11. The host cell of claim 10 wherein said host cell is a bacterial, yeast, insect, plant, viral or mammalian cell.
12. A binding molecule comprising an isolated V_H single domain antibody according to any of claims 1 to 6.
13. The binding molecule according to claim 12 wherein the V_H single domain antibody is linked to a second antigen binding molecule that does not bind to MSLN.

14. The binding molecule according to claim 13 wherein said second antigen binding domain is an antibody or fragment thereof.
15. The binding molecule according to claim 14 wherein said fragment is a Fab, Fab', F(ab')₂, dAb, Fd, Fv, or a single chain Fv fragment, a human heavy chain variable domain (VH), or an isolated CDR.
16. The binding molecule according to any of claims 13 to 15 wherein said antigen is a tumor specific antigen, optionally PSMA.
17. The binding molecule according to any of claims 13 to 14 wherein said antigen is Her2, CD123, CD19, CD20, CD22, CD23, CD74, BCMA, CD30, CD33, CD52, EGFR, CECAM6, CAXII, CD24, CEA, cMet, TAG72, MUC1, MUC16, STEAP, EphvIII, FAP, GD2, IL-13Ra2, L1-CAM, PSCA, GPC3, Her3, gpA33, 5T4 and ROR1, CD3, CDE28, CD27, CD40, GITTA, OX40, CD80, CD86, ICOS.
18. The binding molecule according to any of claims 13 to 15 wherein said antigen is a checkpoint target.
19. The binding molecule according to claim 18 wherein said antigen binding domain binds to an immune checkpoint inhibitor.
20. The binding molecule according to claim 19 wherein said immune checkpoint inhibitor is an anti-PD1, anti PDL-1, anti PDL-2, anti CTL-4, anti -TIM-3 or anti LAG-3 antibody
21. The binding molecule according to any of claims 12 to 20 conjugated to a toxin, enzyme, radioisotope, half-life extending moiety, therapeutic molecule, label or other chemical moiety.
22. The isolated binding molecule according to claim 21 wherein said half-life extending moiety is selected from the group consisting of an albumin binding moiety, a transferrin binding moiety, a polyethylene glycol molecule, a recombinant polyethylene glycol molecule, human serum albumin, a fragment of human serum albumin, and an albumin binding peptide or single domain antibody that binds to human serum albumin.
23. The use of a single domain antibody according to any of claims 1 to 6 in a multispecific or multivalent binding agent.
24. A pharmaceutical composition comprising a V_H single domain antibody according to claim 1 to 6 or a binding molecule according to any of claims 12 to 22 and a pharmaceutical acceptable carrier, excipient or diluent.

25. A method for treating a cancer comprising administering a V_H single domain antibody according to any of claims 1 to 6 or a binding molecule according to any of claims 11 to 22.
26. The V_H single domain antibody according to claim any of claims 1 to 6, a pharmaceutical composition of claim 24 or a binding molecule according to any of claims 12 to 22 for use in the treatment of cancer.
27. The method according to claim 25, a V_H single domain, binding molecule or pharmaceutical composition according to claim 26 wherein the cancer is selected from a haematological cancer or solid cancer.
- 10 28. The method, V_H single domain antibody, binding molecule, pharmaceutical composition of claim 27 wherein the cancer is oesophageal cancer, breast cancer, gastric cancer, cholangiocarcinoma, pancreatic cancer, colon cancer, lung cancer, thymic carcinoma, mesothelioma, ovarian cancer, and endometrial cancer.
29. The method according to any of claims 25, 27 or 28, a V_H single domain antibody, the binding molecule or a pharmaceutical composition according to any of claims 26, 27 or 28 further comprising the step of measuring mesothelin expression in a sample from said subject.
- 20 30. The method according to any of claims 25, 27, 28 or 29, a V_H single domain antibody, the binding molecule or a pharmaceutical composition according to any of claims 26, 27, 28 or 29 wherein the cancer is MSLN-positive.
31. The method according to any of claims 25, 27, 28 to 30, a V_H single domain antibody, the binding molecule or a pharmaceutical composition according to any of claims 26, 27, 28 to 30 further comprising the step of administering an immune checkpoint inhibitor or other therapeutic moiety.
32. The method, the cell, cell population, pharmaceutical composition or use according to claim 31 wherein said immune checkpoint inhibitor is an anti-PD1, anti PDL-1, anti PDL-2, anti CTL-4, anti TIM-3 or anti LAG-3 antibody.
33. An immunoconjugate comprising a single domain antibody according to any of claims 1 to 6 or a binding molecule according to any of claims 12 to 22 linked to a therapeutic agent.
- 30 34. A method for producing a V_H single domain antibody according to any one of claims 1 to 6 comprising expressing a nucleic acid encoding said binding molecule in a host cell and isolating the binding molecule from the host cell.

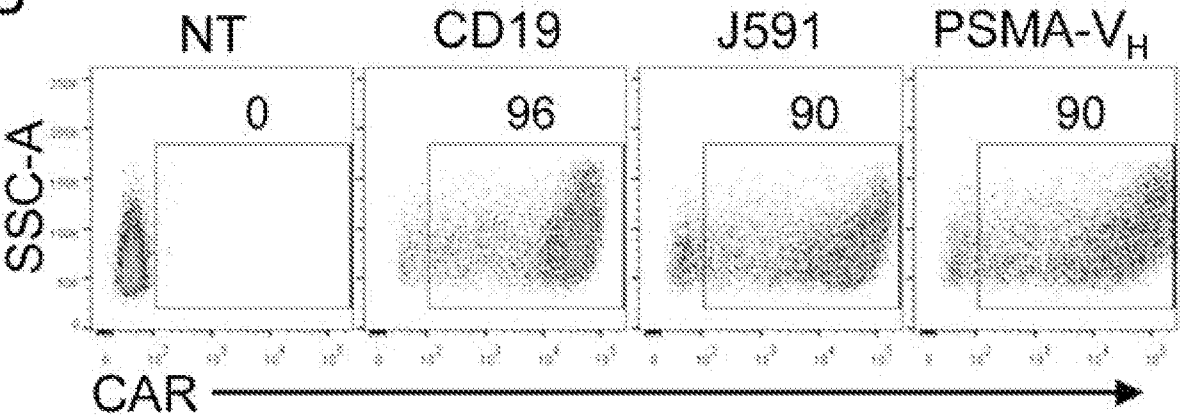
35. A kit comprising a single domain antibody according to any of claims 1 to 6 or a binding molecule according to any of claims 12 to 22, or a pharmaceutical composition according to claim 24 and optionally instructions for use..
36. A method for detecting the presence of human MSLN in a test sample or detecting or diagnosing a cancer in a subject comprising contacting said sample with a single domain antibody according to any of claims 1 to 6 and at least one detectable label and detecting binding of said single domain antibody to human MSLN.
37. The method according to claim 36 wherein the biological sample is a tissue sample or a fluid sample.
- 10 38. A method for producing a single V_H domain antibody that binds to human MSLN comprising
- a) immunising a transgenic animal that expresses a nucleic acid construct comprising human heavy chain V genes and that is not capable of making functional endogenous light or heavy chains with an MSLN antigen,
 - b) generating a library from said animal
 - c) isolating single V_H domain antibodies from said libraries and
 - d) identifying a single V_H domain antibody that binds to human MSLN and
 - e) isolating said antibody.
39. A single V_H domain antibody obtained or obtainable by the method of claim 38.
- 20 40. An isolated heavy chain only antibody comprising a V_H domain that binds to human MSLN.
41. A transgenic rodent that produces a heavy chain only antibody of claim 40.
42. A heavy chain only antibody comprising a V_H domain that binds to human MSLN obtained or obtainable from a transgenic mouse which expresses human V, D and J loci and does not produce functional endogenous lambda and kappa light chains and heavy chain.
43. An antibody drug conjugate comprising a V_H single domain antibody according to any of any of claims 1 to 5 or a binding molecule according to any of claims 11 to 21.

Figure 1

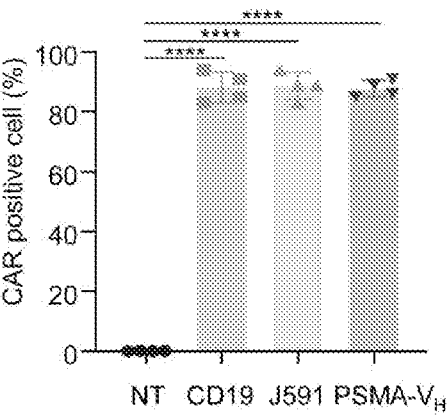
A



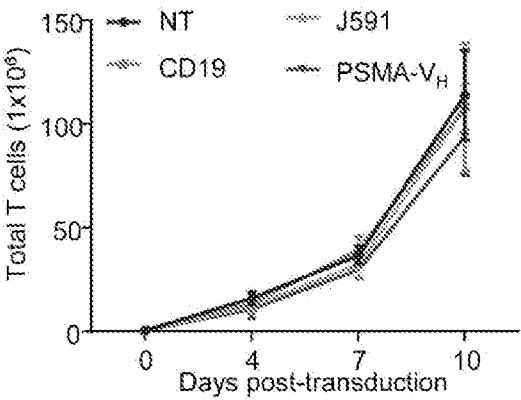
B



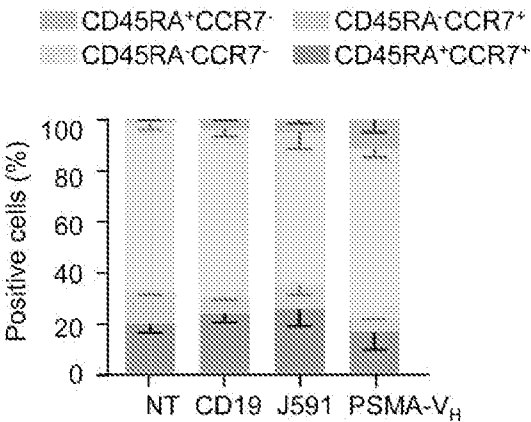
C



D



E



F

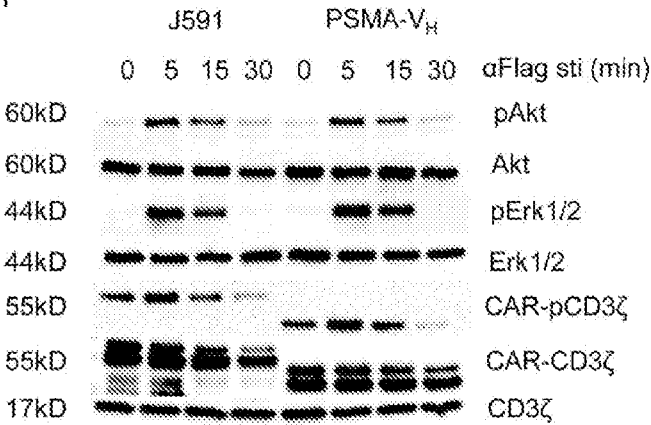
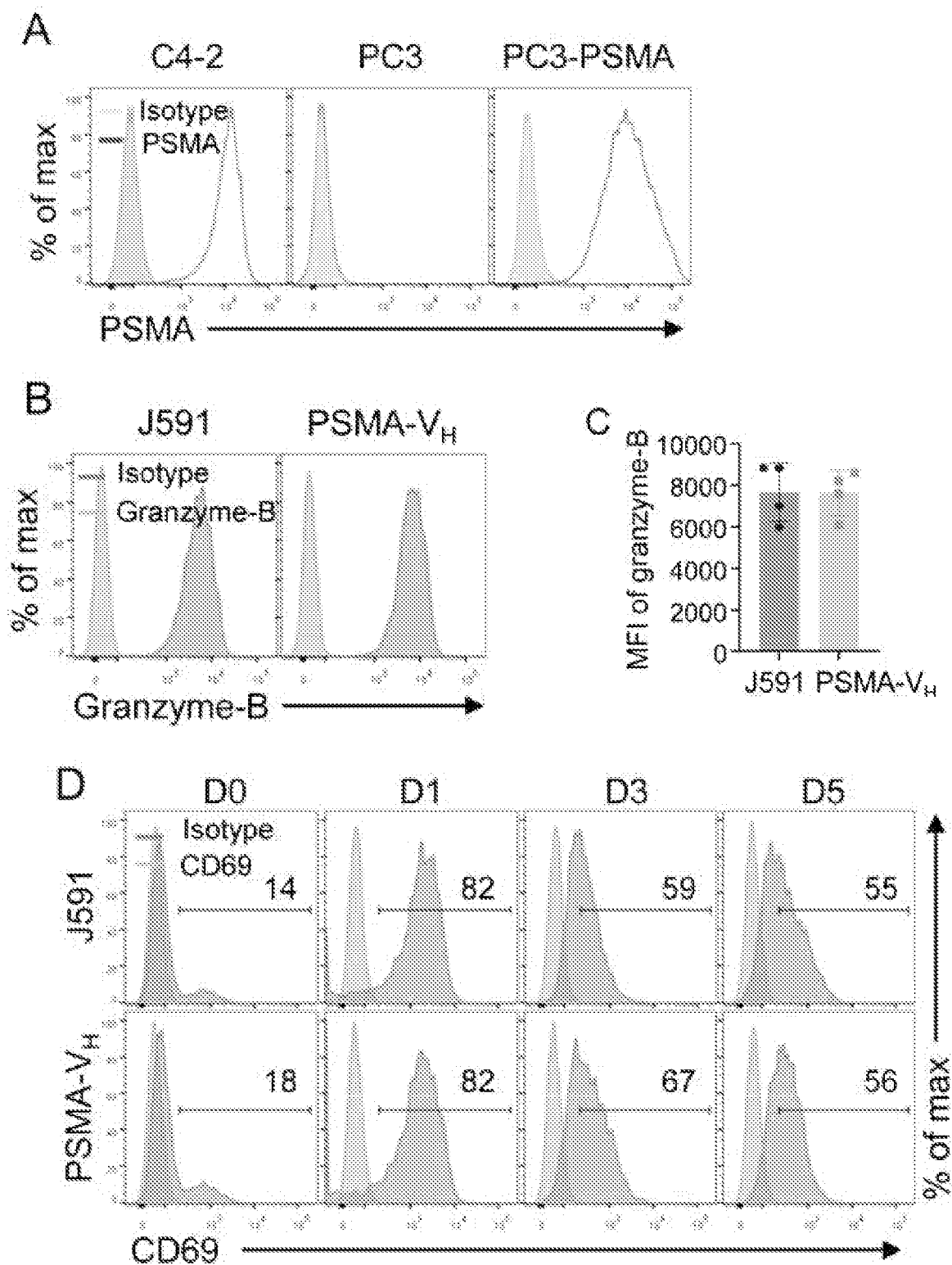


Figure 2



3 / 22

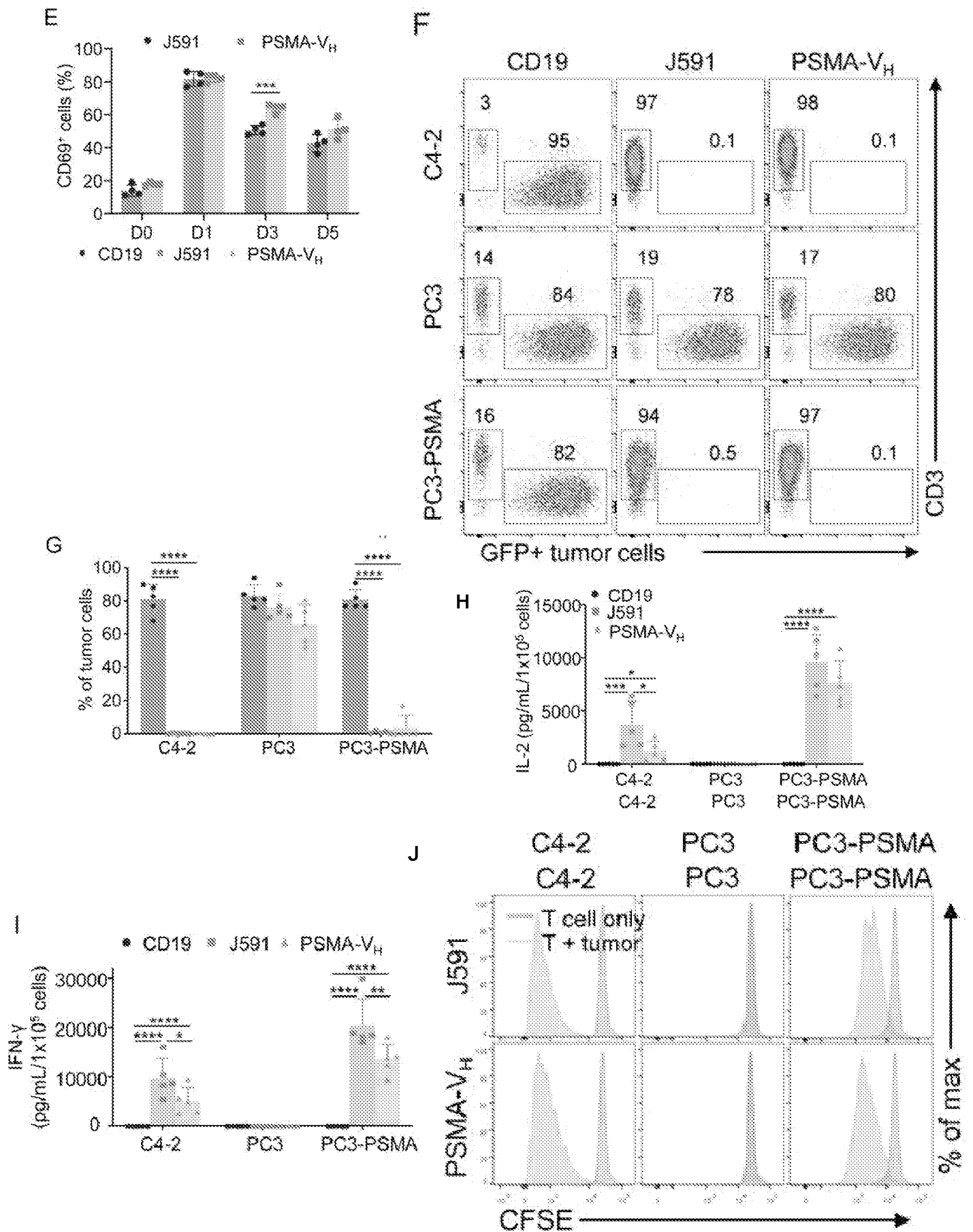
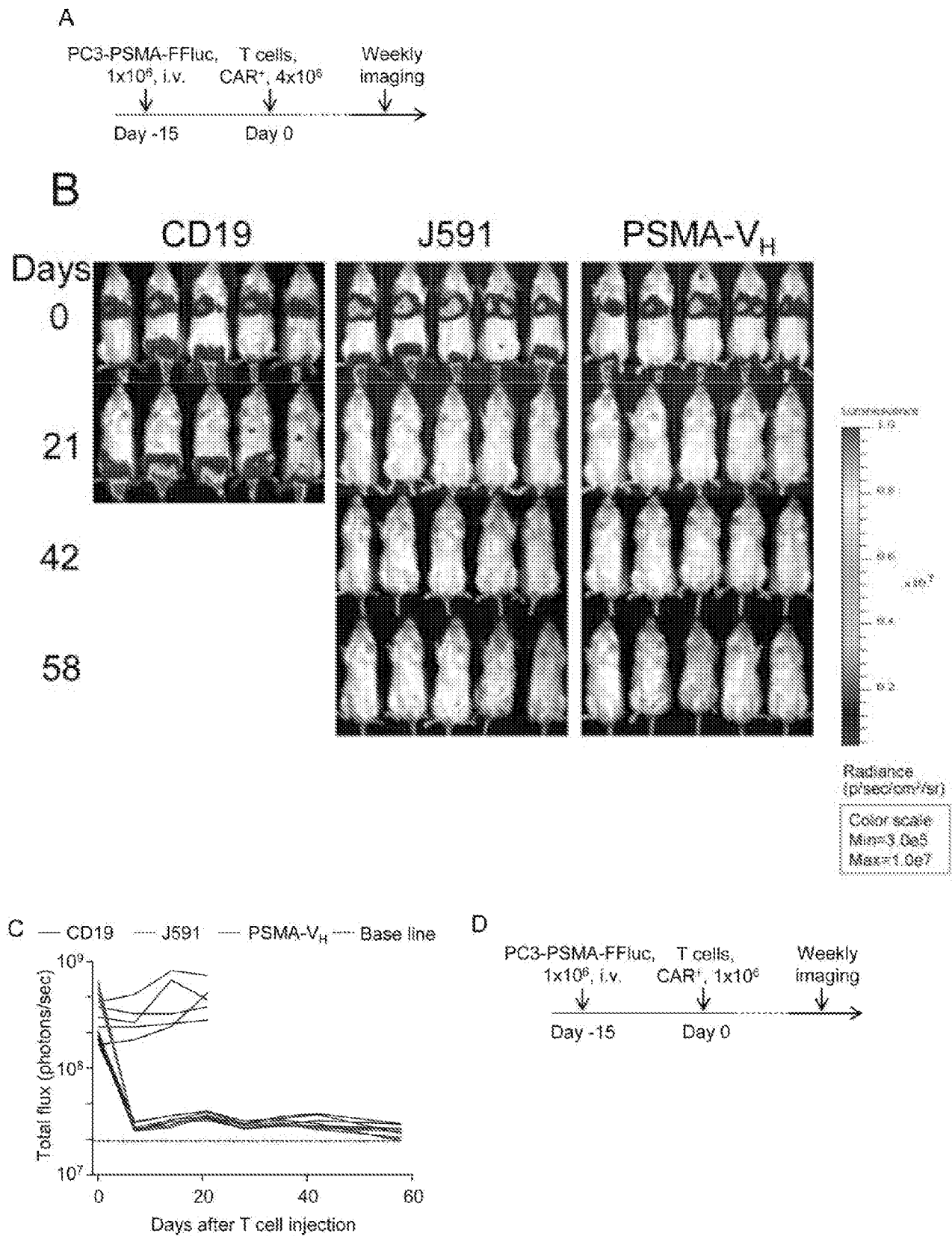
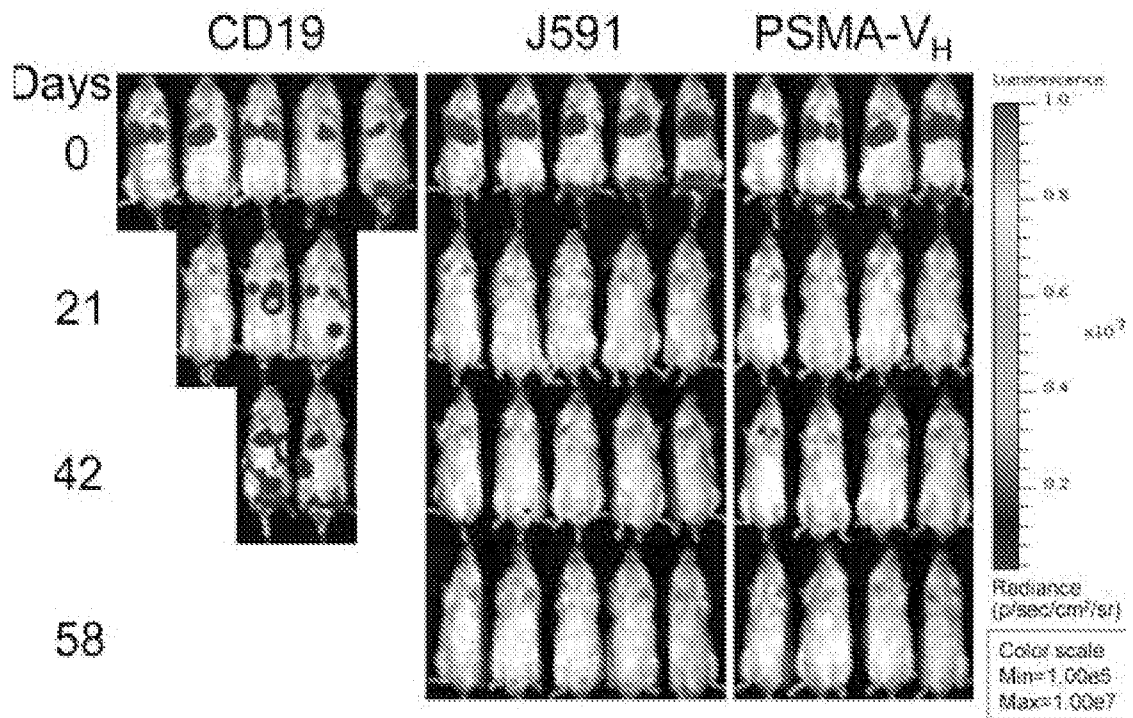
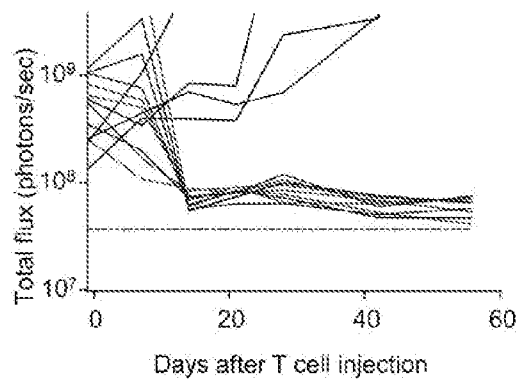


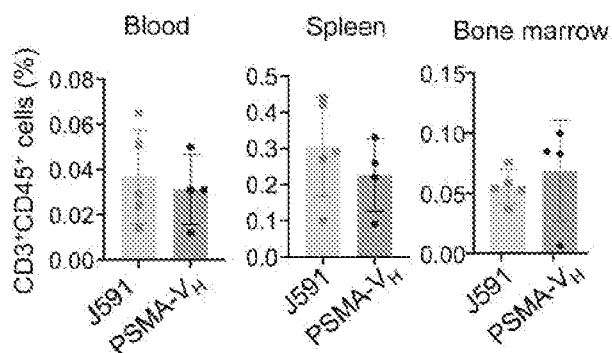
Figure 3



E

F — CD19 — J591 — PSMA-V_H — Base line

G



H

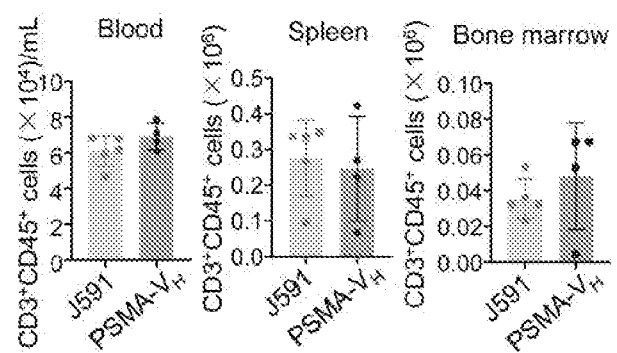
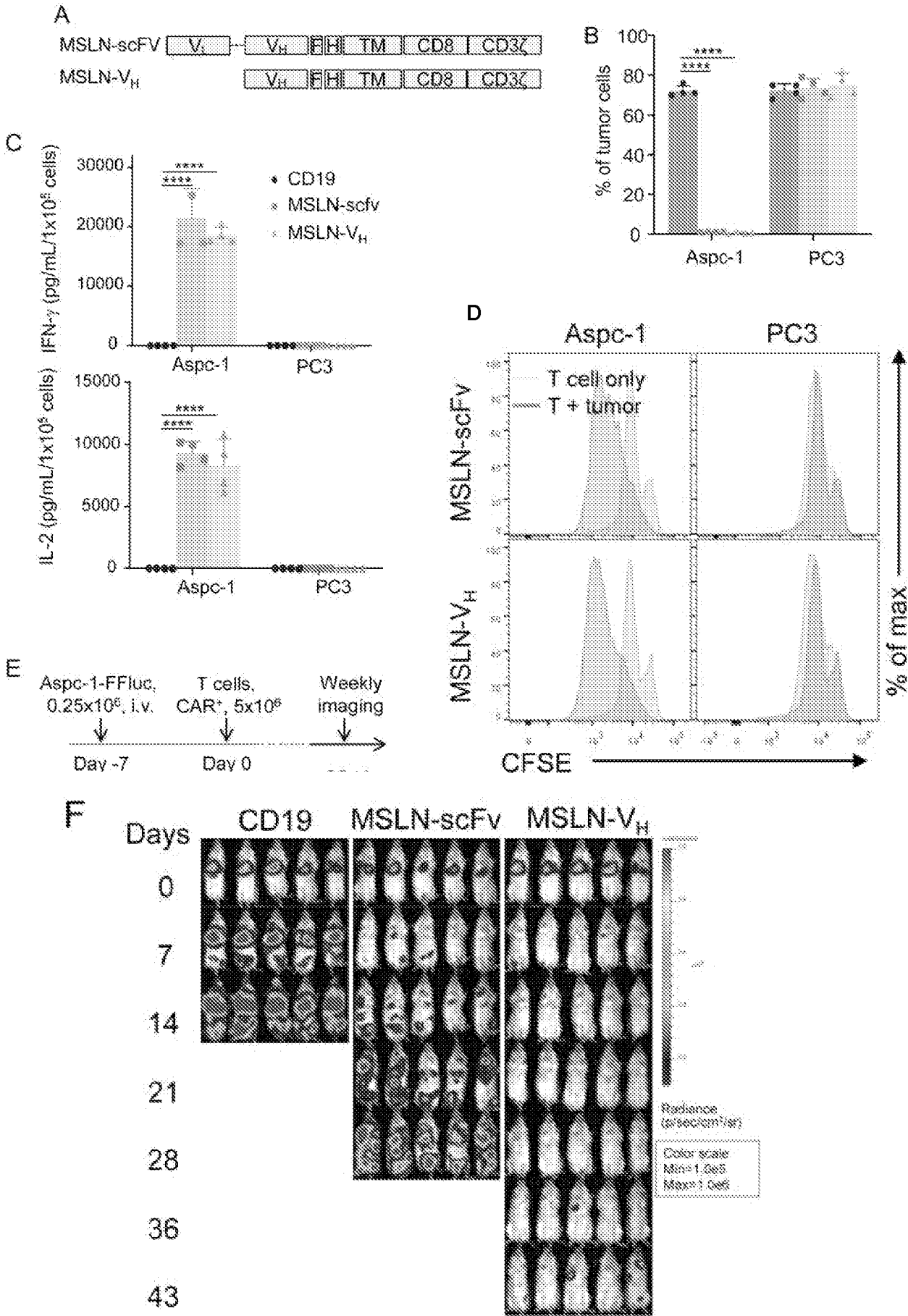


Figure 4



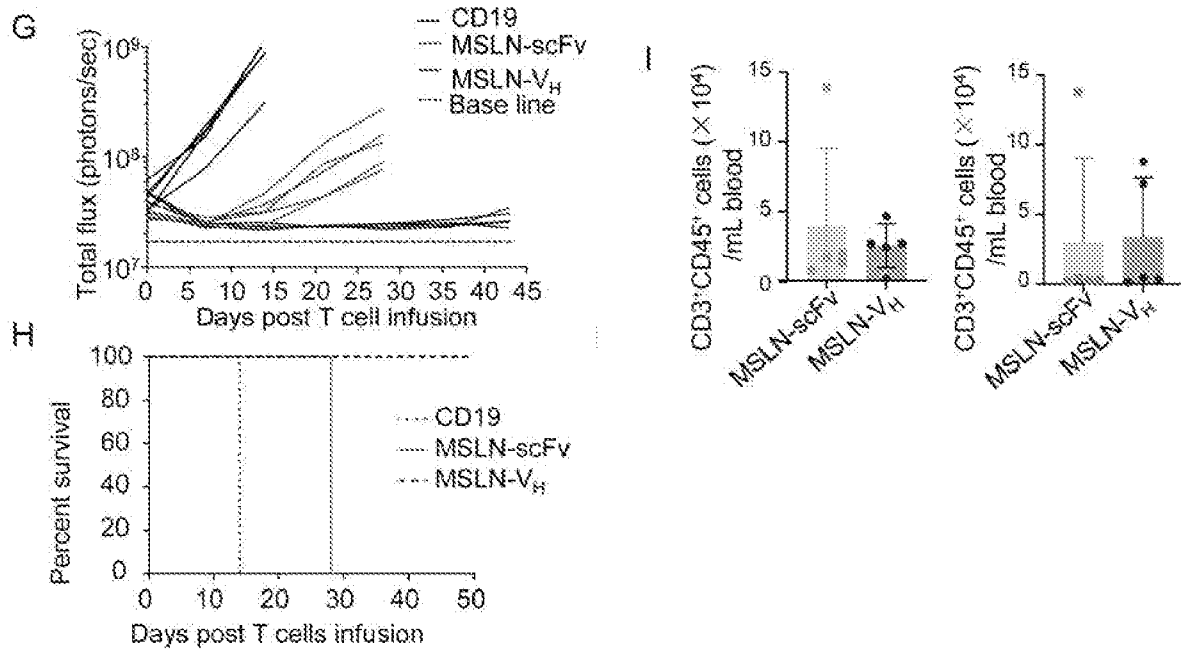


Figure 5

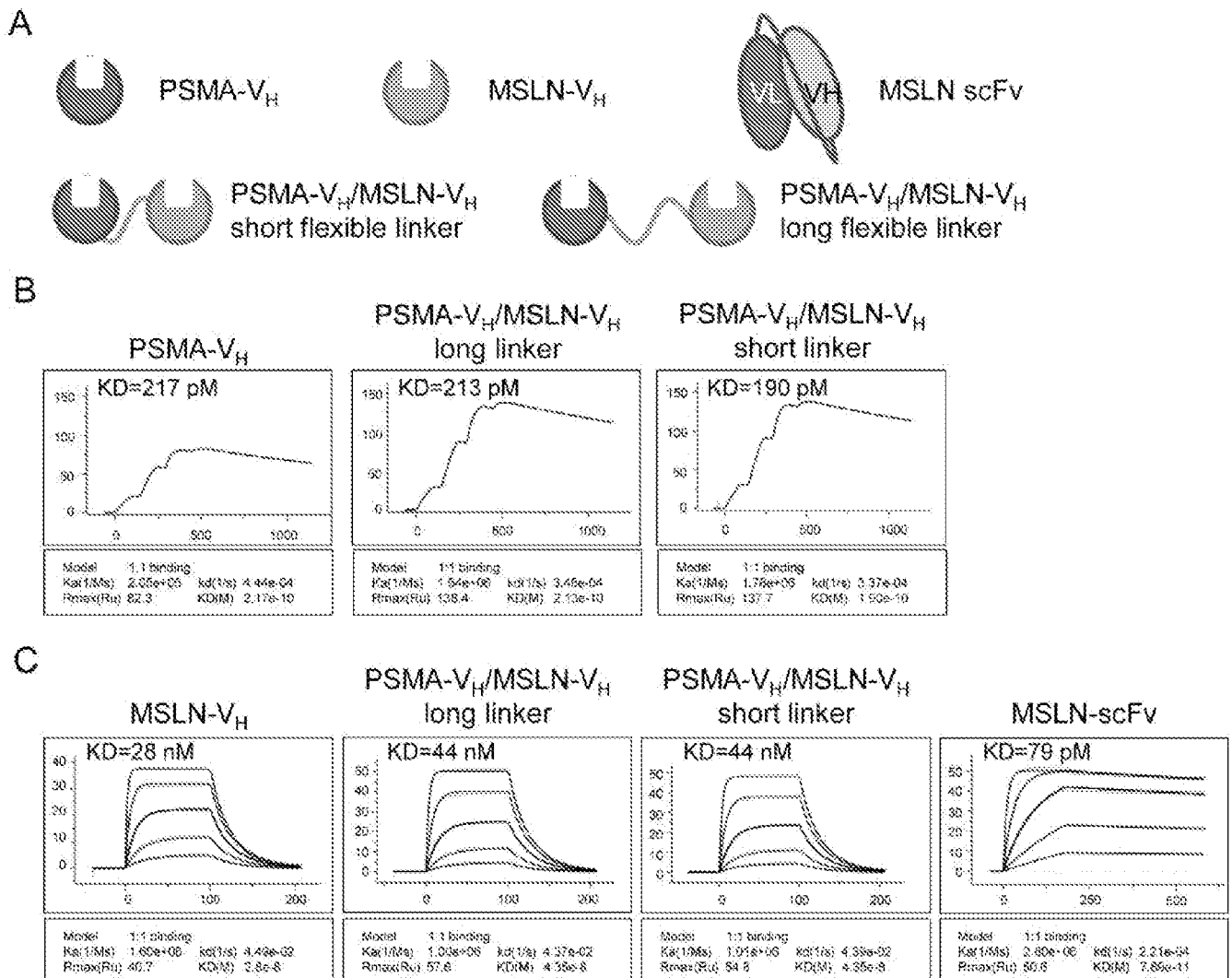


Figure 6

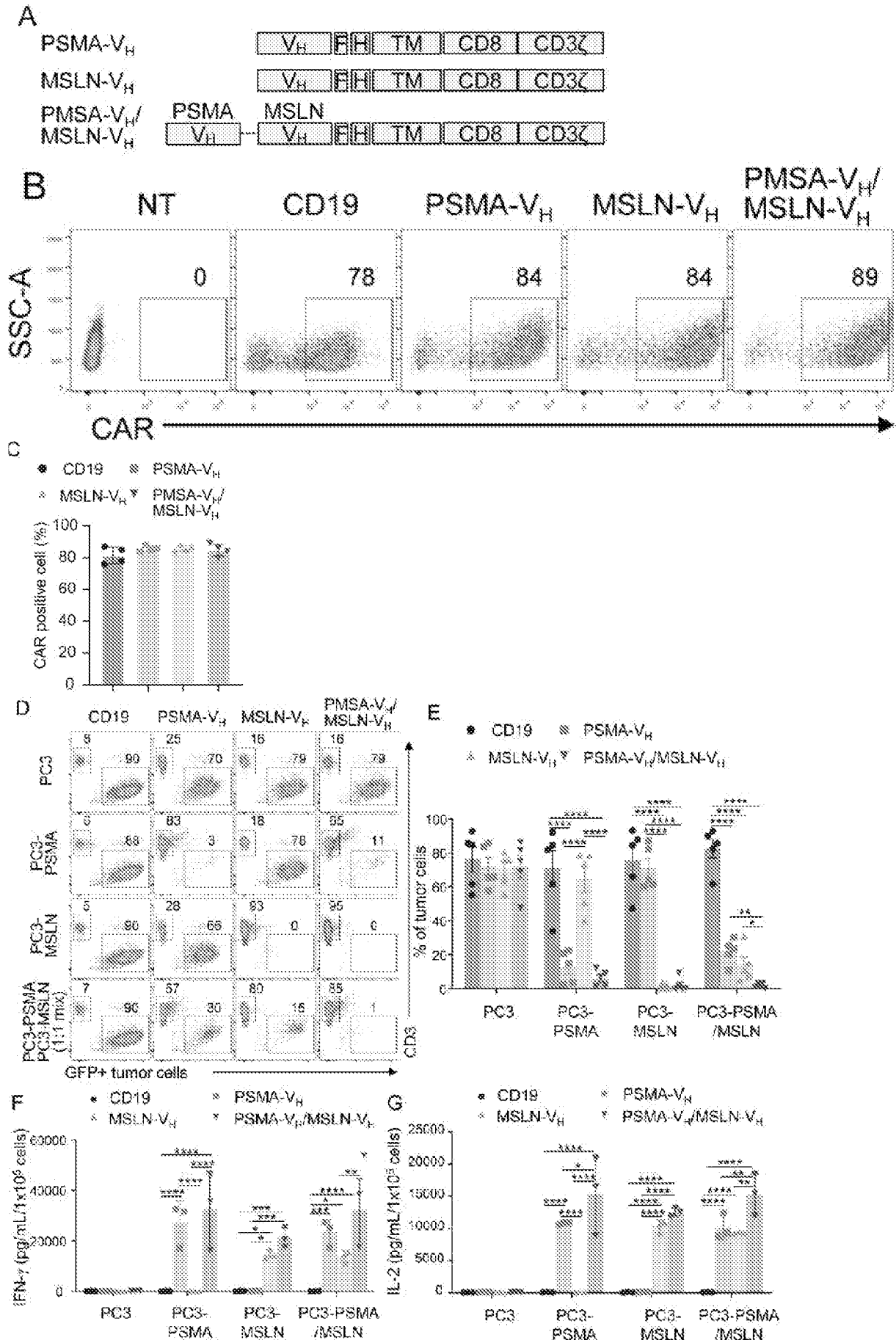


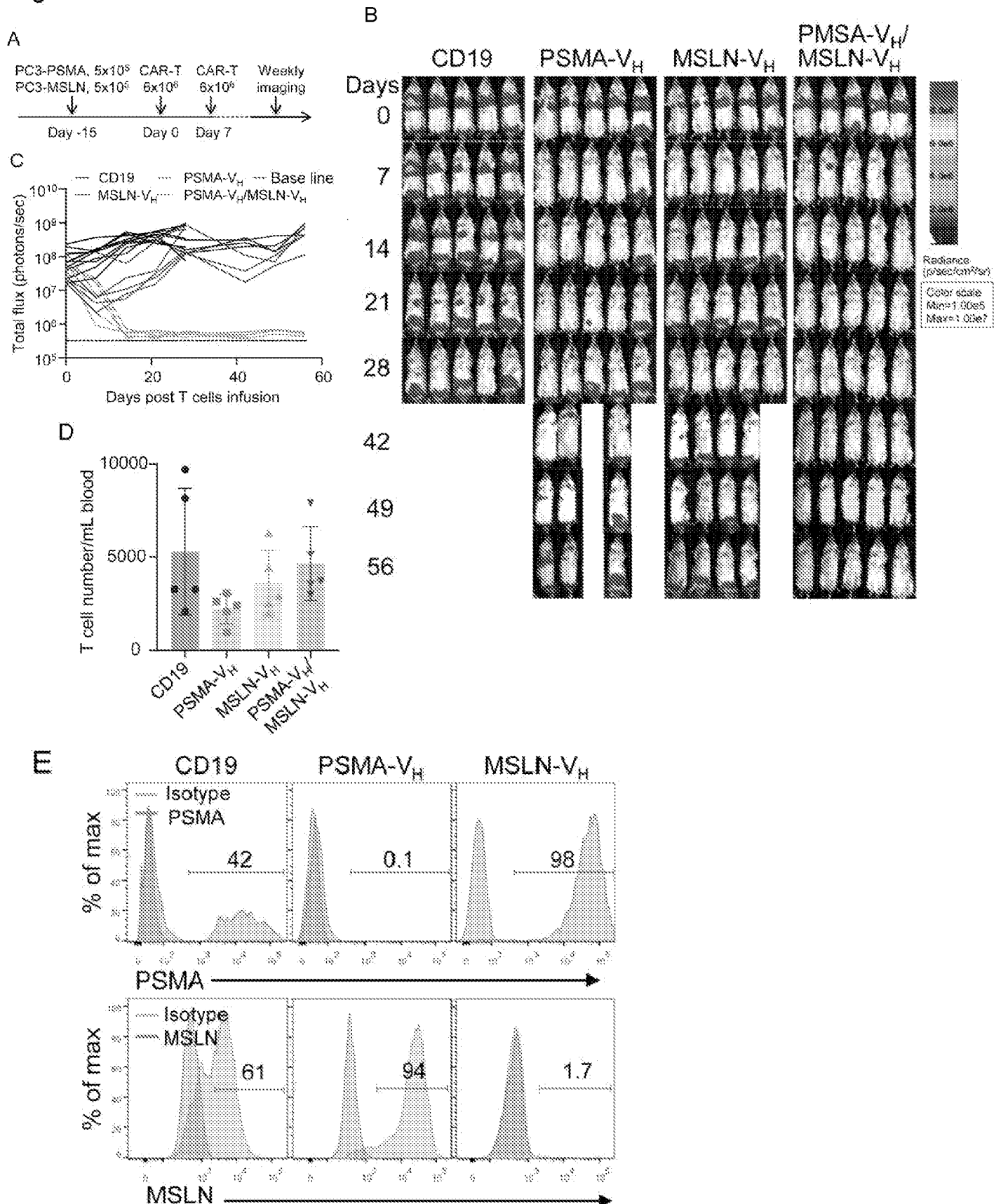
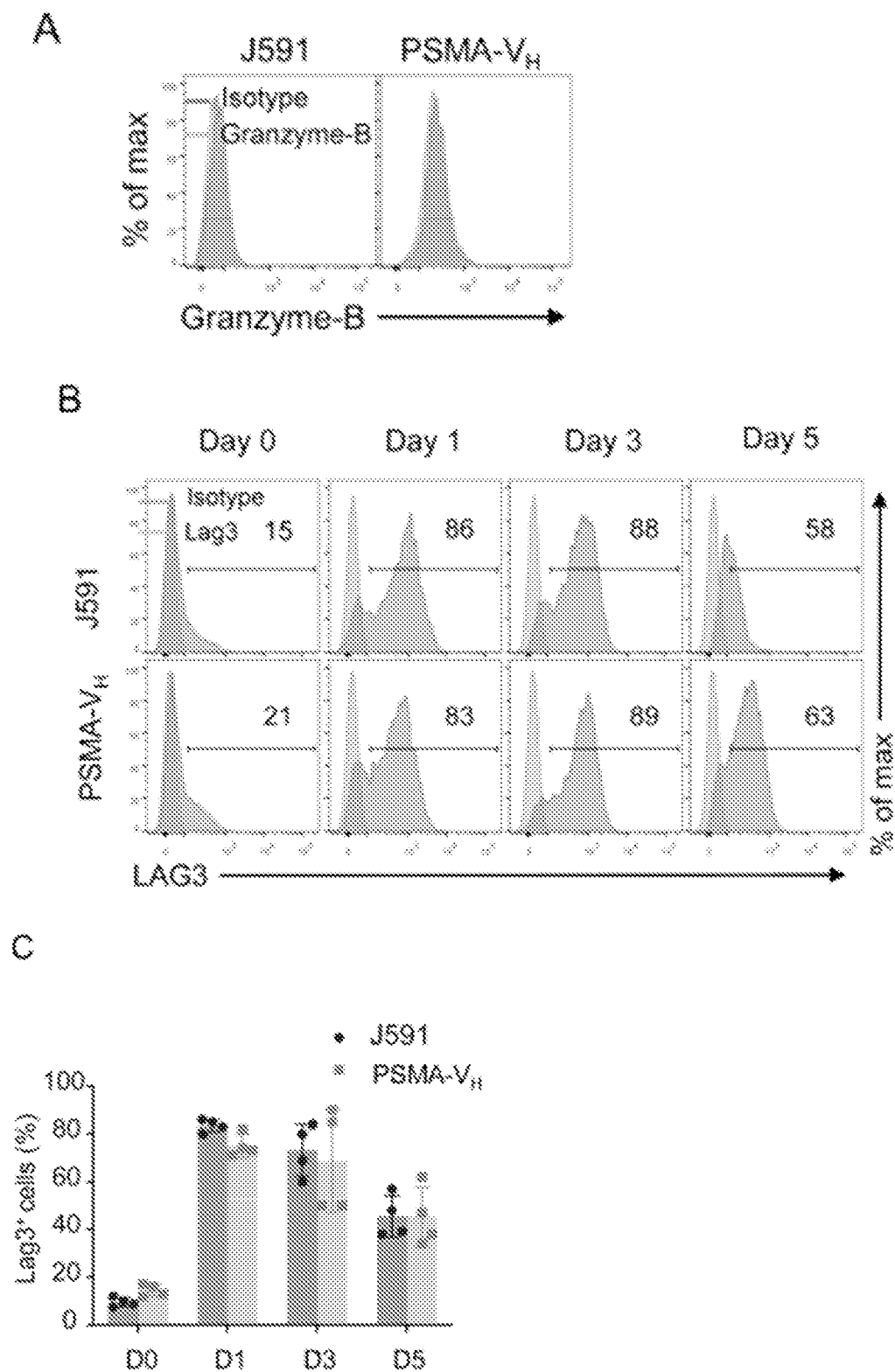
Figure 7

Figure 8



11 / 22

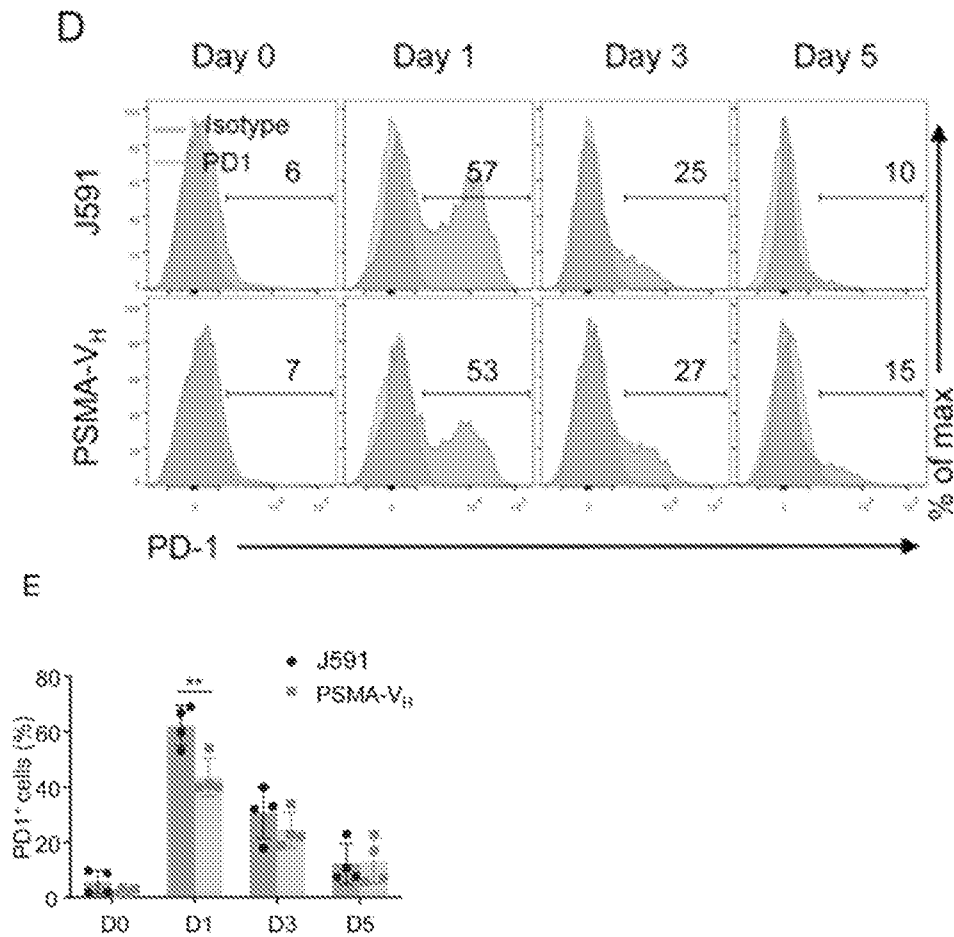


Figure 9

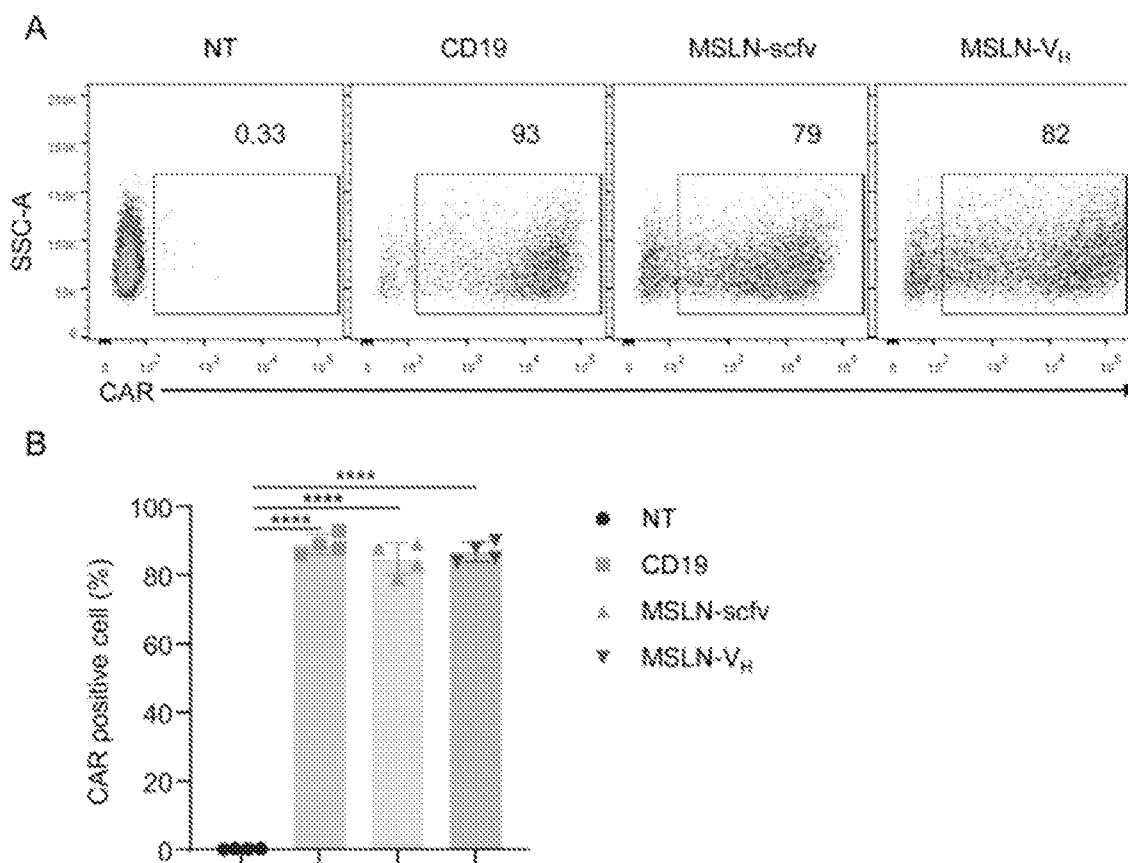


Figure 10

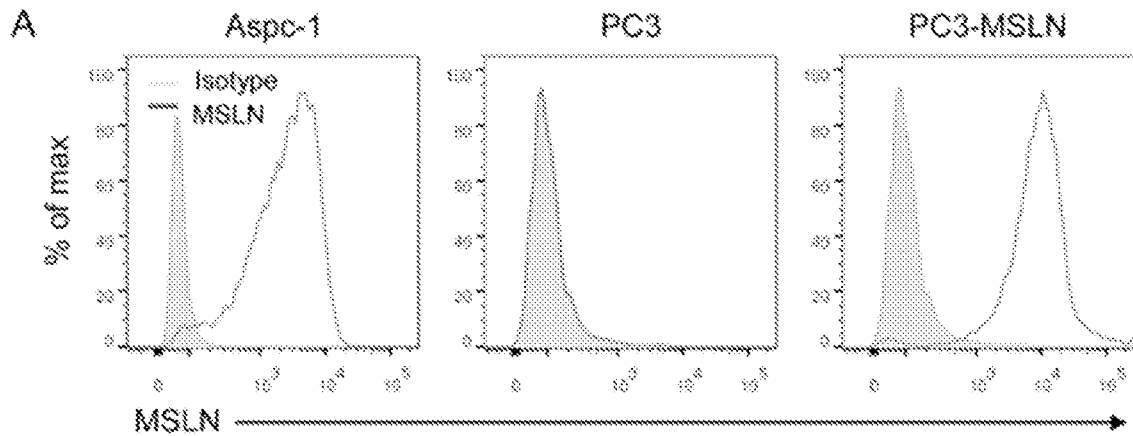
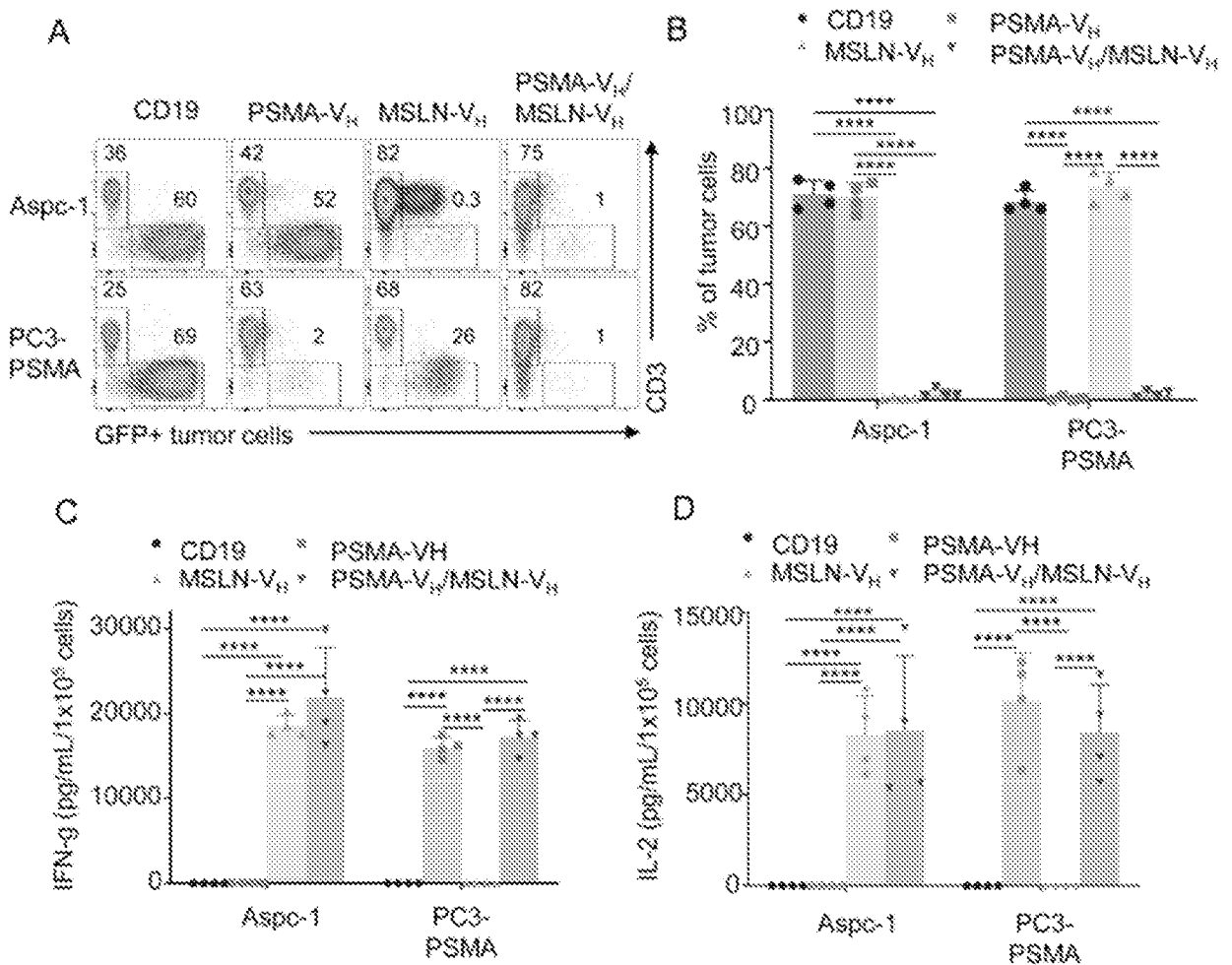


Figure 11



13 / 22

Figure 12

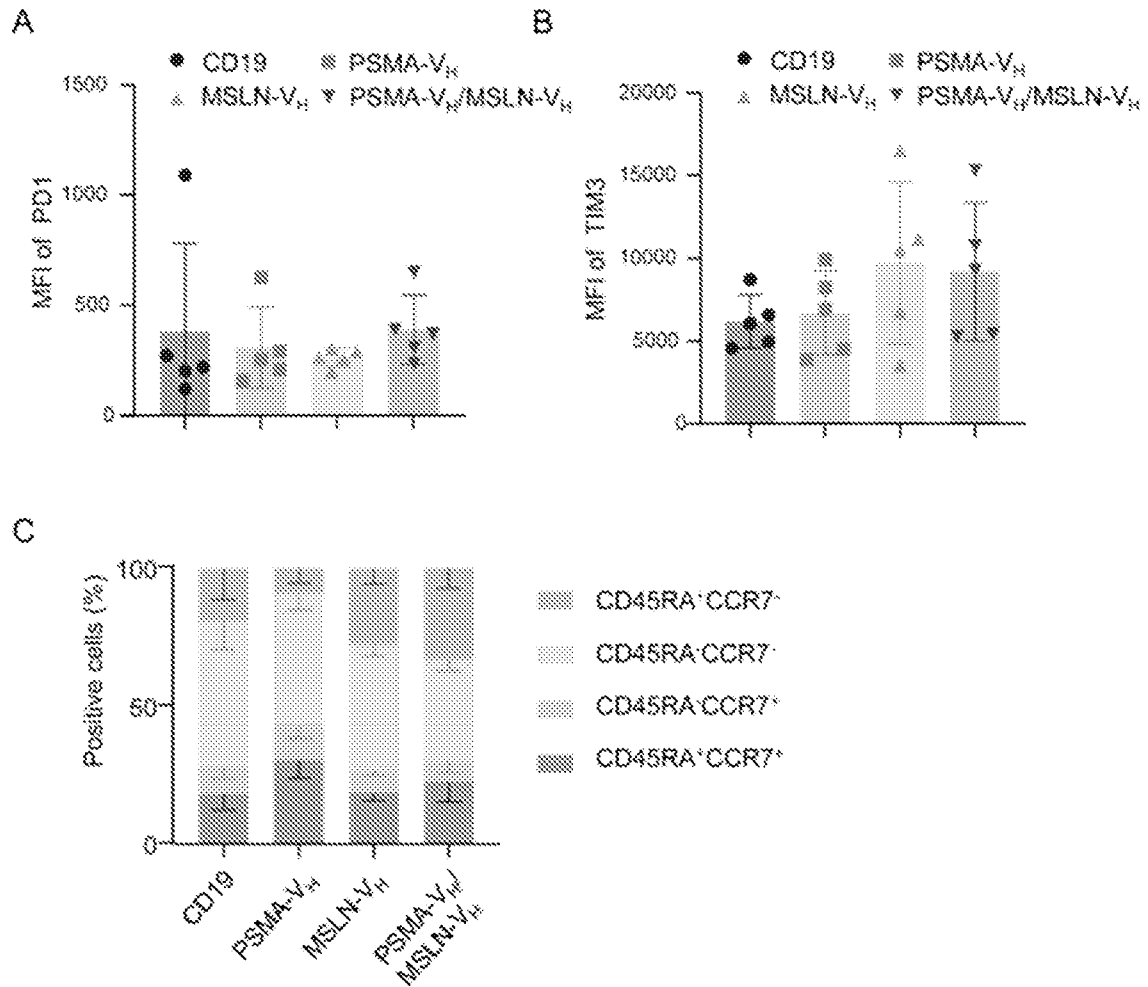


Figure 13 A

1.1

EVQLLESGGGLVQPGGSLRLSCAASGFSFSSYAMSWVRQAPGKGLEWVSSIGENDGTTDYADSVKGRFTISRDN
NSKSTLYLQMNSLRVEDTAVYYCVK**DGVHW**GQGTLTVSS

1.2

EVQLVESGGGLVQPGGSLRLSCAASGFTFSSYAMSWVRQAPGKGLEWVSSIGDNNNSTEYADSVKGRFTISRDN
NSKSTLYLQMNSLSAEDTAVYYCVK**DGVHW**GQGTLTVSS

1.3

EVQLVESGGGLVQPGGSLRLSCAASGFSFSSYAMSWVRQAPGKGLEWVSSIGDNNNSTDYADSVKGRFTISRDN
NSKSTLYLQMNSLRAEDTAVYYCVK**DGVHW**GQGTLTVSS

1.4

EVQLVESGGGLVQPGGSLRLSCAASGFTFSSYAMSWVRQAPGKGLEWVSSIGDGTYYADSVKGRFTISRDN
KSTLYLQMNSLRAEDTAVYYCAK**DGVHW**GQGTLTVSS

1.5

EVQLVESGGGLVQPGGSLRLSCAASGFTFSTYAMSWVRQAPGKGLEWVSSIGENDRTYYVDSVKGRFTISRDN
NSKSTLYLQMNSLRAEDTAVYYCAK**DGVHW**GQGTLTVSS

1.6

QVQLVESGGGLVQPGGSLRLSCAASGFTFSSYAMSWVRQAPGKGLEWVSSIGDNNRTYYADSVKGRFTISRDN
NSKSTLYLQMNSLRAEDTAVYYCAK**DGVHW**GQGTLTVSS

1.7

EVQLVESGGGLVQPGGSLRLSCAASGFTFSSYAMSWVRQAPGKGLEWVSSIGDGTYYADSVKGRFTISRDN
KSTLYLQMNSLRAEDTAVYYCAK**DGVHW**GQGTLTVSS

1.8

EVQLLESGGGLVQPGGSLRLSCAASGFSFSSYAMSWVRQAPGKGLEWVSSIGENDGTTDYADSVKGRFTISRDN
NSKNTLYLQMNSLRVEDTAVYYCVK**DGVHW**GQGTLTVSS

1.9

EVQLLESGGGLVQPGGSLRLSCAASGFSFSSYALSWVRQAPGKGLEWVSSIGENDGTTDYADSVKGRFTISRDN
SKNTLYLQMNSLRVEDTAVYYCVK**DGVHW**GQGTLTVSS

1.10

EVQLLESGGGLVQPGGSLRLSCAASGFSFSSYALSWVRQAPGKGLEWVSSIGENNATTDYADFVKGRFTISRDN
SKNTLYLQMNSLRVEDTAVYYCVK**DGVHW**GQGTLTVSS

1.11

EVQLLESGGGLVQPGGSLRLSCAASGFSFSSYALSWVRQAPGKGLEWVSSIGENNDDTDYADNVKGRFTISRDN
NSKNTLYLQMNSLRVEDTAVYYCVK**DGVHW**GQGTLTVSS

1.12

EVQLLESGGGLVQPGGSLRLSCAASGFSFSSYALSWVRQAPGKGLEWVSSIGENNATTDYADAVKGRFTISRDN
SKNTLYLQMNSLRVEDTAVYYCVK**DGVHW**GQGTLTVSS

1.13

EVQLLESGGGLVQPGGSLRLSCAASGFSFSSYALSWVRQAPGKGLEWVSSIGENNHTTDYAADVKGRTISRDN
SKNTLYLQMNSLRVEDTAVYYCVK**DGVHW**GQGTLTVSS

1.14

EVQLLESGGGLVQPGGSLRLSCAASGFSFSSYALSWVRQAPGKGLEWVSSIGENNATTDYADVKGRTISRDN
SKNTLYLQMNSLRVEDTAVYYCVK**DGVHW**GQGTLTVSS

1.15

EVQLLESGGGLVQPGGSLRLSCAASGFSFSSYALSWVRQAPGKGLEWVSSIGENNHTTDYAADFVKGRFTISRDN
SKNTLYLQMNSLRVEDTAVYYCVK**DGVHW**GQGTLTVSS

1.16

EVQLLESGGGLVQPGGSLRLSCAASGFSFSS**SYALS**WVRQAPGKGLEWVSS**SIGENNHTTDYADTVKGRFTISRDN**
SKNTLYLQMNSLRVEDTAVYYCVK**DGVH**WGQGTLLTVSS

1.17

EVQLLESGGGLVQPGGSLRLSCAASGFSFSS**SYALS**WVRQAPGKGLEWVSS**SIGENNDDTDYADAVKGRFTISRDN**
SKNTLYLQMNSLRVEDTAVYYCVK**DGVH**WGQGTLLTVSS

1.18

EVQLLESGGGLVQPGGSLRLSCAASGFSFSS**SYALS**WVRQAPGKGLEWVSS**SIGENNATTDYAASVKGRFTISRDN**
SKNTLYLQMNSLRVEDTAVYYCVK**DGVH**WGQGTLLTVSS

1.19

EVQLLESGGGLVQPGGSLRLSCAASGFSFSS**SYALS**WVRQAPGKGLEWVSS**SIGENNDDTDYAAYVKGRFTISRDN**
SKNTLYLQMNSLRVEDTAVYYCVK**DGVH**WGQGTLLTVSS

1.20

EVQLLESGGGLVQPGGSLRLSCAASGFSFSS**SYALS**WVRQAPGKGLEWVSS**SIGENNHTTDYAATVKGRFTISRDN**
SKNTLYLQMNSLRVEDTAVYYCVK**DGVH**WGQGTLLTVSS

Figure 13B

2.1

EVQLVESGGGVVQPGRSLRLSCAASGFSFSG**YGMH**WVRQAPGKGLEWV**AYISYDGSNKYYADSVKGRFTISRDN**
SKNTLYLQMNSLRAEDTAVYYCAK**DPAWGLRLGESSYDFDI**WGQGTMTVSS

2.3

EVQLVESGGGVVQPGRSLRLSCAASGFSFSG**YGMH**WVRQAPGKGLEWV**AHISYDGSNRYAESVKGRFTISREN**
SKNTLSLQMNSLRAEDTAVYYCAK**DPAWGLRLGELSSYDFDI**WGQGTMTVSS

2.4

QVTLKESGGGVVQPGRSLKLSCAASGFSFSG**YGMH**WVRQAPGKGLEWV**AVISYDGSNRYADSVKGRFTISRDN**
SKNTLSLQMNSLRAEDTAVYYCARD**DPAWGLRLGELSSYDFEI**WGQGTMTVSS

2.5

QVQLVESGGGVVQPGRSLRLSCAASGFSFSG**YGMH**WVRQAPGKGLEWV**AVISYDGSNRYADSVKGRFTISRD**
NSKNTLSLQMNSLRAEDTAVYYCAK**DPAWGLRLGELSSYDFEI**WGQGTMTVSS

2.6

EVQLVESGGGVVQPGRSLRLSCAASGFSFSG**YGMH**WVRQAPGKGLEWV**AVISYDGSNKYYADSVKGRFTISRDN**
SKNTLYLQMNSLRAEDTAVYYCAK**DPAWGLRLGELSSYKFEI**WGQGTMTVSS

2.7

EVQLVESGGGVVQPGRSLRLSCAASGFSFSG**YGMH**WVRQAPGKGLEWV**ALISYDGSNKYYADSVKGRFTISRDN**
SKNTLYLQMNSLRAEDTAVYYCAK**DPAWGLRLGEQSSYAFDI**WGQGTMTVSS

2.8

QVQLVESGGGVVQPGRSLRLSCAASGFSFSG**YGMH**WVRQAPGKGLEWV**VISYDGSNKYYADSVKGRFTISRDN**
SKNTLYLQMNSLRTEDTAVYYCAK**DPAWGLRLGEQSSYAFEI**WGQGTMTVSS

2.9

EVQLLESGGGVVQPGRSLRLSCAASGFSFSG**YGMH**WVRQAPGKGLEWV**AVISYDGSNKYYADSVKGRFTISRDN**
SKNTLYLQMNSLRVEDTAVYYCAK**DPAWGLRLGEQSSYAFEI**RQGQGTMTVSS

2.10

EVQLVESGGGVVQPGRSLRLSCAASGFTFS**GYGMH**WVRQAPGKGLEWVAYISYDGSNRYYADSVKGRFTISRDN
SKKTLISLQMNSLRAEDTAVYYCAK**DPAWGLRLGESSYDFDI**WGQGTMTVTVSS

2.11

EVQLVESGGGVVQPGRSLRLSCAASGFSFS**GYGLH**WVRQAPGKGLEWVAYISYDESNKYYAPSVKGRFTISRDNS
KNTLYLQMNSLRAEDTAVYYCAK**DPAWGLRLGESSYDFDI**WGQGTMTVTVSS

2.12

EVQLVESGGGVVQPGRSLRLSCAASGFSFS**GYGMH**WVRQAPGKGLEWVAYISYDKNKYYADKVKGRFTISRDN
SKNTLYLQMNSLRAEDTAVYYCAK**DPAWGLRLGESSYDFDI**WGQGTMTVTVSS

2.13

EVQLVESGGGVVQPGRSLRLSCAASGFSFS**GYGLH**WVRQAPGKGLEWVAYISYDASNKYYADNVKGRFTISRDNS
KNTLYLQMNSLRAEDTAVYYCAK**DPAWGLRLGESSYDFDI**WGQGTMTVTVSS

2.14

EVQLVESGGGVVQPGRSLRLSCAASGFSFS**GYGVH**WVRQAPGKGLEWVAYISYDASNKYYADNVKGRFTISRDN
SKNTLYLQMNSLRAEDTAVYYCAK**DPAWGLRLGESSYDFDI**WGQGTMTVTVSS

2.15

EVQLVESGGGVVQPGRSLRLSCAASGFSFS**GYGLH**WVRQAPGKGLEWVAYISYDKNKYYADKVKGRFTISRDNS
KNTLYLQMNSLRAEDTAVYYCAK**DPAWGLRLGESSYDFDI**WGQGTMTVTVSS

2.16

EVQLVESGGGVVQPGRSLRLSCAASGFSFS**GYGAH**WVRQAPGKGLEWVAYISYDKNKYYADKVKGRFTISRDNS
KNTLYLQMNSLRAEDTAVYYCAK**DPAWGLRLGESSYDFDI**WGQGTMTVTVSS

2.17

EVQLVESGGGVVQPGRSLRLSCAASGFSFS**GYGMH**WVRQAPGKGLEWVAYISYDASNKYYADNVKGRFTISRDN
SKNTLYLQMNSLRAEDTAVYYCAK**DPAWGLRLGESSYDFDI**WGQGTMTVTVSS

2.18

EVQLVESGGGVVQPGRSLRLSCAASGFSFS**GYGQH**WVRQAPGKGLEWVAYISYDASNKYYADNVKGRFTISRDN
SKNTLYLQMNSLRAEDTAVYYCAK**DPAWGLRLGESSYDFDI**WGQGTMTVTVSS

2.19

EVQLVESGGGVVQPGRSLRLSCAASGFSFS**GYGFH**WVRQAPGKGLEWVAYISYDASNKYYADNVKGRFTISRDNS
KNTLYLQMNSLRAEDTAVYYCAK**DPAWGLRLGESSYDFDI**WGQGTMTVTVSS

2.20

EVQLVESGGGVVQPGRSLRLSCAASGFSFS**GYGMH**WVRQAPGKGLEWVAIISYDGSNRYYADSVKGRFTISRDN
SKNTLSLQMNSLRAEDTAVYYCAK**DPAWGLRLGESSYDFEI**WGQGTMTVTVSS

2.21

QVQLVESGGGVVQPGRSLKLSCAASGFSFS**GYGMH**WVRQAPGKGLEWVAVISYDGSNRYYADSVKGRFTISRD
NSKNTLSLQMNSLRAEDTAVYYCAK**DPAWGLRLGKLSSYDFEI**WGQGTMTVTVSS

2.22

EVQLVESGGGVVQPGRSLRLSCAASGFSFS**GYGTH**WVRQAPGKGLEWVAYISYDGSNKYYAAPVKGRFTISRDNS
KNTLYLQMNSLRAEDTAVYYCAK**DAAWGLRLGESSYDFDI**WGQGTMTVTVSS

2.23

EVQLVESGGGVVQPGRSLRLSCAASGFSFS**GYGTHWVRQAPGKGLEWVAYISYDESNKYYASSVKGRFTISRDNS**
KNTLYLQMNSLRAEDTAVYYCAK**DRAWGLRLGESSYDFDIWGQGTMTVSS**

2.24

EVQLVESGGGVVQPGRSLRLSCAASGFSFS**GYGMHWVRQAPGKGLEWVAYISYDESNKYYARLVKGRFTISRDNS**
KNTLYLQMNSLRAEDTAVYYCAK**DTAWGLRLGESSYDFDIWGQGTMTVSS**

2.25

EVQLVESGGGVVQPGRSLRLSCAASGFSFS**GYGLHWVRQAPGKGLEWVAYISYDLSNKYYARGVKGRFTISRDNS**
KNTLYLQMNSLRAEDTAVYYCAK**DVAWGLRLGESSYDFDIWGQGTMTVSS**

Figure 13C

3.1

EVQLVESGGGVVQPGRSLRLSCAASGFPLI**SYGMHWVRQAPGKGLEWVAFMTYDGSNRYYADSVKGRFTISRD**
NSKNTLYLQMNSLRDEDTALYYCARD**DRIVGGRVPDAFDIWGQGTMTVSS**

3.2

EVQLVESGGGVVQPGRSLRLSCAASGFPLI**SYGMNWVRQAPGKGLDWVAFISYDGSNKYYADSVKGRFTISKDNS**
KNTLYLQMNSLRAEDTAVYYCAK**DRIVGARVPDAFDIWGQGTMTVSS**

3.3

EVQLVESGGGVVQPGRSLRLSCAASGFPLI**SYGMNWVRQAPGKGLEWVAFISYDGSNRYYADSVKGRFTISRDNS**
KNTLYLQMNSLRAEDTAVYYCAK**DRIVGARVPDAFDIWGQGTMTVSS**

3.4

EVQLVESGGGAVQPGRSLRLSCAASGFPLI**SYGMNWVRQAPGKGLDWVAFITYDGSNRYYADSVKGRFTISRDN**
SKNTLYLQMNSLRPEDTAVYYCAK**DRIVGARVPDAYDIWGQGTMTVSS**

3.5

QVQLVESGGGVVQPGRSLRLSCAASGFPLI**SYGMNWVRQAPGKGLDWVAFITYDGSNRYYADSVKGRFTISRDN**
SKNTLYLQMNSLRPEDTAVYYCAK**DRIVGARVPDAYDIWGQGTMTVSS**

3.6

EVQLLESGGGVVQPGRSLRLSCAASGFPLI**SYGMNWVRQAPGKGLDWVAFITYDGSNRYYADSVKGRFTISRDNS**
KNTLYLQMNSLRPEDTAVYYCAK**DRIVGARVPDAYDIWGQGTMTVSS**

3.7

QVQLVESGGGLVQPGGSLRLSCAASGFPLI**SYGMNWVRQAPGKGLDWVAFITYDGSNRYYADSVKGRFTISRDN**
SKNTLHLQMDSLRPEDTAVYYCAK**DRIVGARVPDAYDIWGQGTMTVSS**

3.8

EVQLVESGGGVVQPGRSLRLSCAASGFPLI**SYGMNWVRQAPGKGLDWVAFITYDGSNRYYADSVKGRFTISRDN**
SKNTLYLQMNSLRPEDTAVYYCAK**DRIVGARVPDAYDIWGQGLTVSS**

3.9

QVQLVESGGGVVQPGRSLRLSCAASGFPLISYGMNWVRQAPGKGLEWVAFISYDGSNRYYADSVKGRFTISRDN
SKNTLYLQMNSLRPEDTAVYYCAKDRIVGARVPDAYDIWGQGTMTVSS

3.10

QVQLVESGGGVVQPGRSLRLSCAASGFPLISYGMNWVRQAPGKGLDWVAFITYDGSNRYYADSVKGRFTISRDN
SKNTLHLQMNSLRPEDTAVYYCAKDRIVGARVPDAYDIWGQGTMTVSS

3.11

EVQLVESGGGVVQPGRSLRLSCAASGFPLISYGMNWVRQAPGKGLDWVAFITYDGSNRYYADSVKGRFTISRDN
SKNTLHLQMNSLRPEDTAVYYCAKDRIVGARVPDAYDIWGQGTMTVSS

3.12

EVQLVESGGGVVQPGRSLRLSCAASGFPLISYGMNWVRQAPGKGLDWVAFITYDGSNRYYADSVKGRFTISRDN
SKNTLYLQMNSLRPEDTAVYYCAKDRIVGARVPDAYDIWGQGTMTVSS

3.13

QVQLVESGGGVVQPGRSLRLSCAASGFPLISYGMNWVRQAPGKGLDWVAFITYDGSNRYYADSVKGRFTISRDN
SKNTLYLQMNSLRPEDTAVYYCAKDRIVGARVPDAYDIWGQGTMTVSS

3.14

EVQLVESGGGVVRPGGSLRLSCAASGFPLISYGMNWVRQAPGKGLDWVAFITYDGSNRYYADSVKGRFTISRDN
KNTLHLQMNSLRPEDTAVYYCAKDRIVGARVPDAYDIWGQGTMTVSS

3.15

EVQLVESGGGLVQPGGSLRLSCAASGFPLISYGMNWVRQAPGKGLDWVAFITYDGSNRYYADSVKGRFTISRDN
KNTLHLQMNSLRPEDTAVYYCAKDRIVGARVPDAYDIWGQGTMTVSS

3.16

EVQLLESGGGVVQPGRSLRLSCAASGFPLISYGMNWVRQAPGKGLDWVAFITYDGSNRYYADSVKGRFTISRDN
KNTLHLQMNSLRPEDTAVYYCAKDRIVGARVPDAYDIWGQGTMTVSS

3.17

EVQLLESGGGVVQPGRSLRLSCAASGFPLISYGMNWVRQAPGKGLDWVAFITYDGSNRYYADSVKGRFTISRDN
KNTLYLQMNSLKPEDTAVYYCAKDRIVGARVPDAYDIWGQGTMTVSS

3.18

EVQLVESGGGVVQPGRSLRLSCAASGFPLISYGMNWVRQAPGKGLDWVAFITYDGSNRYYADSVKGRFTISRDN
SKNTLYLQMNSLKPEDTAVYYCAKDRIVGARVPDAYDIWGQGTMTVSS

3.19

QVQLVESGGGVVQPGRSLRLSCAASGFPLISYGMHWVRQAPGKGLEWVAFMTYDGSNRYYADAVKGRFTISRDN
NSKNTLYLQMNSLRAEDTAVYYCARDIVGGRVPDAFDIWGQGTMTVSS

3.20

QVQLVESGGGVVQPGRSLRLSCAASGFPLISYGMHWVRQAPGKGLEWVAFQTYDGSNRYYADAVKGRFTISRDN
NSKNTLYLQMNSLRAEDTAVYYCARDRIVGGRVPDAFDIWGQGTMTVSS

3.21

QVQLVESGGGVVQPGRSLRLSCAASGFPLISYGMHWVRQAPGKGLEWVAFQTYDGSNRYYADSVKGRFTISRDN
NSKNTLYLQMNSLRAEDTAVYYCARDRIVGGRVPDAFDIWGQGTMTVSS

3.22

QVQLVESGGGVVQPGRSLRLSCAASGFPLISYGMHWVRQAPGKGLEWVAFQTYDASNRYYADSVKGRFTISRDN
SKNTLYLQMNSLRAEDTAVYYCARDRIVGGRVPDAFDIWGQGTMTVSS

3.23

QVQLVESGGGVVQPGRSLRLSCAASGFPLISYGMHWVRQAPGKGLEWVAFQTYDASNRYYADAVKGRFTISRDN
NSKNTLYLQMNSLRAEDTAVYYCARDRIVGGRVPDAFDIWGQGTMTVSS

3.24

QVQLVESGGGVVQPGRSLRLSCAASGFPLISYGMNHWVRQAPGKGLEWVAFITYDGSNRYYADSVKGRFTISRDN
SKNTLYLQMNSLRAEDTAVYYCAKDRIVGARVPDAYDIWGQGTMTVSS

Figure 13D

4.1

QVQLVESGGGVVQPGRSLRLSCVASGFPLISYGMHWVRQAPGKGREWVAVISYDGSNRYYADSVKGRFTISRDN
SKNTLYLQMNSLRPEDTAVYYCAKERIFGVLTDPDDFDIWGQGTMTVSS

4.2

QVQLVESGGGVVQPGRSLRLSCAASGFPLISYGMHWVRQAPGKGLEWVAVISYDGSNRYYADSVKGRFTISRDN
SKNTLYLQMNSLRPEDTAVYYCAKERIFGVLTDPDDFDIWGQGTMTVSS

4.3

EVQLLESGGGVVQPGRSLRLSCAASGFPLISYGMHWVRQAPGKGLEWVAVISYDGANRYYADSVKGRFTISRDN
SKNTLYLQMNSLRPEDTAVYYCAKERIFGVLTDPDDFEIWGQGTMTVSS

4.4

EVQLVESGGGVVQPGRSLRLSCAASGFTFTSYGMHWVRQAPGKGLEWVAVISYDGSNRYYADSVKGRFTISRDN
SKNTLYLQMNSLRPEDTAVYYCAKERIFGALTPDDFDIWGQGTMTVSS

Figure 13E

5.1

QVQLVESGGGVVQPGRSLRLSCAASGFTFNHYGMHWVRQAPGKGLEWVAIISYDGNKYYTDSVKGRFTISRDN
SKNTLYLQMNSLRVEDTAVYYCAKGLWPSDVWGQGTMTVSS

5.2

EVQLVESGGGVVQPGRSLRLSCAASGFTFNHYGMHWVRQAPGKGLEWVAIISYDGNKYYTDSVKGRFTISRDN
SKNTLYLQMNSLRVEDTAVYYCAKGLWPSDVWGQGTMTVSS

Figure 13F

6.1

QVQLQESGPGLVKPSQTLSTCTVSGGSIS**NSGYYS**WVRQHPGKDLEWIG**FIYYNGSIHYNPSL**KSRVVISVDTSK
NQFSLKMNSVTAADTAVYYCARD**DGDDYGDY**LRGQGT LTVSS

6.2

QVQLQESGPGLVKPSQTLSTCTVSGGSIS**NSGYYS**WIRQHPGKGLEWIG**FIYYNGSIHYNPSL**KSRVVISVDTSKN
QFSLKMSSVTAADTAVYYCARD**DGDDYGDY**LRGQGT LTVSS

6.3

QVQLQESGPGLVKPSQTLSTCTVSGGSIS**NSGYYS**WVRQHPGKGLEWIG**FIYYNGSIHYNPSL**KSRVVISVDTSK
NQFSLKLNSVTAADTAVYYCARD**DGDDYGDY**LRGQGT LTVSS

6.4

QVQLQESGPGLVKPSQTLSTCTVSGGSIS**NSGYYS**WIRQHPGKGLEWIG**FIYYNGSIHYNPSL**KSRVVISVDTSKN
QFSLKLSSVTAADTAVYYCARD**DGDDYGDY**LRGQGT LTVSS

6.5

QVQLQESGPGLVKPSQTLSTCTVSGGSIS**NSGYYS**WVRQHPGKGLEWIG**FIYYNGSIHYNPSL**KSRVTISVDTSK
NQFSLKMSSVTAADTAVYYCARD**DGDDYGDY**LRGQGT LTVSS

6.6

QVQLQESGPGLVKPSQTLSTCTVSGGSIS**NSGYYS**WVRQHPGKGLEWIG**FIYYNGSIHYNPSL**KSRVTISVDTSK
NQFSLKLNSVTAADTAVYYCARD**DGDDYGDY**LRGQGT LTVSS

6.7

QVQLQESGPGLVKPSQTLSTCTVSGGSIS**NSGYYS**WVRQHPGKGLEWIG**FIYYNGSIHYNPSL**KSRVTISVDTSK
NQFSLKLSSVTAADTAVYYCARD**DGDDYGDY**LRGQGT LTVSS

Figure 13G

7.1

EVQLVESGGGLVQPGGSLRLSCAASGFTF**SSYWMY**WVRQAPGKGLEWVANINHDGSEKYY**VDSVKGR**FTISR
NAKNSLYLQMNSLRAEDTAVYYCARD**DSLIVGERGY**WGQGT LTVSS

7.2

EVQLVESGGGLVQPGGSLRLSCAASGFTF**SSYWMY**WVRQAPGKGLEWVANINHDGSEKYY**VDSVKGR**FTISR
NAKNSLYLQMNSLRAEDTAVYYCARD**DNLIVGERGY**WGQGT LTVSS

7.3

EVQLVESGGGLVQPGGSLRLSCAASGFTF**SSYWMY**WVRQAPGKGLEWVANINHDGSEKYY**VDSVKGR**FTISR
NAKNSLYLQMNSLRAEDTAVYYCARD**DSLIVGERGY**WGQGT LTVSS

7.4

EVQLVESGGGLVQPGGSLRLSCAASGFTFSS**SYWMY**WVRQAPGKGLEWVANIN**HQGSEKYYVDSVKGRFTISRD**
NAKNSLYLQMNSLRAEDTAVYYCARD**SLIVGERGYWGQ**GLTVTVSS

7.5

EVQLVESGGGLVQPGGSLRLSCAASGFTFSS**SYWMY**WVRQAPGKGLEWVANIN**HPGSEKYYVDSVKGRFTISRDN**
AKNSLYLQMNSLRAEDTAVYYCARD**SLIVGERGYWGQ**GLTVTVSS

7.6

EVQLVESGGGLVQPGGSLRLSCAASGFTFSS**SYWMY**WVRQAPGKGLEWVANIN**HEGSEKYYVDSVKGRFTISRDN**
AKNSLYLQMNSLRAEDTAVYYCARD**SLIVGERGYWGQ**GLTVTVSS

7.7

EVQLVESGGGLVQPGGSLRLSCAASGFTFSS**SYWMY**WVRQAPGKGLEWVANIN**HIGSEKYYVDSVKGRFTISRDN**
AKNSLYLQMNSLRAEDTAVYYCARD**SLIVGERGYWGQ**GLTVTVSS

7.8

EVQLVESGGGLVQPGGSLRLSCAASGFTFSS**SYWMY**WVRQAPGKGLEWVANIN**HGDSEKYYVDSVKGRFTISRD**
NAKNSLYLQMNSLRAEDTAVYYCARD**TLIVGERGYWGQ**GLTVTVSS

Figure 13H

8.1

QVQLQQWGAGLLKPSETLSLTCAVYGGSFSS**GYWWS**WIRQPPGKGLEWIGEIN**HSGSTNYPNPSLKSRVTISVDTSK**
NQFSLKLSSVTAADTAVYYCARG**PIPATAIPDAFDI**WGQGTMTVTVSS

Figure 13I

9.1

QVQLQQWGAGLLKPSETLSLTCAVYGGSFSS**GHYWS**WIRQPPGKGLEWIGDIN**HSGSTNYPNPSLKSRVTISVDTSK**
NQFSLKLSSVTAADTAVYYCVRD**YGDSRSLFDY**WGQGTMTVTVSS

Figure 13J

10.1

QVQLVESGGGLVQPGGSLRLSCAASGFTFSS**SYGMH**WVRQAPGKGLEWVAF**MSYDGSNKKYYVDSVKGRFTISRD**
NSKNTLYLQMNSLRAEDTAVYYCAK**GDYDFWSGYPDYDM**DVWGQGTMTVTVSS

Figure 13K

11.1

EVQLVESGGGLVKPGGSLRLSCAASGFNLS**SYGM**YWVRQAPGKGLEWVAV**ISYDGSNKNYADSVKGRFTISRDNS**
KNTLFLQMNSLRVEDTAVYYCAK**GGNALYSSGWPDDG**FDIRGQGTMTVTVSS

Figure 13L

QVQLVESGGGVVQGRSLRLSCAASGFTFS**NFGMH**WARQAPGKGLEWVAV**ISYDGSNKKYYADTVKGRFTISRD**
NSKNTLYLEMNSLRADDTAVYYCAK**GLWPPMDVRGQ**GTMTVTVSS

22 / 22

Figure 13M

13.1

EVQLVESGGGSVQPGGSLRLSCAASGFTFS**DYWMT**WVRQVPGKGLEWVAN**IKQDGSEKYYVDSVKGR**FTISRD
NAKNSLYLQMNSLRAEDTAVYYC**ARDRGG**AV**ALYHNGMDM**GGQGTTTVTVSS

Figure 13N

14.1

KCSWWSLGEVKKPGASVKVSCASGYTFT**SYDIN**WVRQATGQGLEWMG**WMNP**NSGNTGY**AQKFQGR**VTMT
RNTSISTAYMELSSLRSED**TAVYYCARGNGPGITGTTD**YWGQGLTVTVSS

Figure 13O

15.1

EVQLVESGGGVVRPGGSLRLSCAASGFTFD**DYGMS**WVRQAPGKGLEWVSG**INWNGDRTGYADSVKGR**FTISRD
NAKNSLYLQMNSLRAEDTALYYCGRE**NVIVPAATY**WGQGLTVTVSS

INTERNATIONAL SEARCH REPORT

International application No

PCT/GB2023/050997

A. CLASSIFICATION OF SUBJECT MATTER

INV. C07K16/28 C07K16/30 C12N15/52 C12N15/62 A61P35/00
 C07K14/435 C07K14/725 C07K14/705

ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C07K C12N A61P

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

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, BIOSIS, EMBASE, SCISEARCH, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WANG GUANMENG ET AL: "Fully human antibody V H domains to generate mono and bispecific CAR to target solid tumors", JOURNAL FOR IMMUNOTHERAPY OF CANCER, vol. 9, no. 4, 1 April 2021 (2021-04-01), page e002173, XP093054299, DOI: 10.1136/jitc-2020-002173 figures 1-4, 6, 7 pages 2-3 -----	1-43
Y	WO 2017/191476 A1 (CRESCENDO BIOLOGICS LTD [GB]) 9 November 2017 (2017-11-09) examples 1-4; sequence 24 -----	13-37
Y	WO 2019/012260 A1 (CRESCENDO BIOLOGICS LTD [GB]) 17 January 2019 (2019-01-17) examples 5-12 ----- -/-	13-37



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents :

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search

14 June 2023

Date of mailing of the international search report

26/06/2023

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2

NL - 2280 HV Rijswijk

Tel. (+31-70) 340-2040,

Fax: (+31-70) 340-3016

Authorized officer

Domingues, Helena

INTERNATIONAL SEARCH REPORT

International application No

PCT/GB2023/050997

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 2018/127709 A1 (CRESCENDO BIOLOGICS LTD [GB]) 12 July 2018 (2018-07-12) examples 1-11 -----	13-37
Y	WO 2020/229842 A1 (CRESCENDO BIOLOGICS LTD [GB]) 19 November 2020 (2020-11-19) examples 1-8 -----	13-37
E	WO 2023/064876 A1 (TENEOBIO INC [US]; ARSENAL BIOSCIENCES INC [US]) 20 April 2023 (2023-04-20) examples 1-5 -----	1-43

INTERNATIONAL SEARCH REPORT

International application No.

PCT/GB2023/050997

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

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:
 - a. ☒ forming part of the international application as filed.
 - b. ☐ furnished subsequent to the international filing date for the purposes of international search (Rule 13~~ter~~1 (a)).

☐ accompanied by a statement to the effect that the sequence listing does not go beyond the disclosure in the international application as filed.
2. ☐ With regard to any nucleotide and/or amino acid sequence disclosed in the international application, this report has been established to the extent that a meaningful search could be carried out without a WIPO Standard ST.26 compliant sequence listing.
3. Additional comments:

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/GB2023/050997

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2017191476 A1	09-11-2017	CN 109641948 A	16-04-2019
		CN 116162174 A	26-05-2023
		EP 3452503 A1	13-03-2019
		US 2019144561 A1	16-05-2019
		WO 2017191476 A1	09-11-2017
WO 2019012260 A1	17-01-2019	CN 111051348 A	21-04-2020
		EP 3652213 A1	20-05-2020
		JP 2020530285 A	22-10-2020
		US 2020131274 A1	30-04-2020
		WO 2019012260 A1	17-01-2019
WO 2018127709 A1	12-07-2018	CN 110121510 A	13-08-2019
		CN 110167966 A	23-08-2019
		EP 3565840 A1	13-11-2019
		EP 3565841 A1	13-11-2019
		EP 3565842 A1	13-11-2019
		JP 7183163 B2	05-12-2022
		JP 7267921 B2	02-05-2023
		JP 2020503870 A	06-02-2020
		JP 2020515235 A	28-05-2020
		JP 2022191316 A	27-12-2022
		JP 2023010753 A	20-01-2023
		US 2019322749 A1	24-10-2019
		US 2020239570 A1	30-07-2020
		US 2020239573 A1	30-07-2020
		US 2022306744 A1	29-09-2022
		WO 2018127709 A1	12-07-2018
		WO 2018127710 A1	12-07-2018
		WO 2018127711 A1	12-07-2018
WO 2020229842 A1	19-11-2020	AU 2020275926 A1	18-11-2021
		CA 3136598 A1	19-11-2021
		CN 113840835 A	24-12-2021
		EP 3969473 A1	23-03-2022
		EP 3969474 A1	23-03-2022
		IL 287555 A	01-12-2021
		JP 2022532370 A	14-07-2022
		KR 20220008839 A	21-01-2022
		SG 11202111671Q A	29-11-2021
		US 2022220215 A1	14-07-2022
		US 2022227850 A1	21-07-2022
		WO 2020229842 A1	19-11-2020
		WO 2020229844 A1	19-11-2020
WO 2023064876 A1	20-04-2023	NONE	