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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
23 June 2022 (23.06.2022)

WIPO | PCT

(10) International Publication Number
WO 2022/133089 A1

(51) International Patent Classification:
C07K 16/28 (2006.01) *A61P 35/00* (2006.01)

(21) International Application Number:
PCT/US2021/063812

(22) International Filing Date:
16 December 2021 (16.12.2021)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
63/127,885 18 December 2020 (18.12.2020) US
63/262,312 08 October 2021 (08.10.2021) US

(71) Applicant: **BIOARDIS, LLC** [US/US]; 7310 Miramar Road, San Diego, California 92126 (US).

(72) Inventors: **HAERIZADEH, Farzad**; 7310 Miramar Road, San Diego, California 92126 (US). **TAYEBI, Masood**; 7310 Miramar Road, San Diego, California 92126 (US).

(74) Agent: **FULLER, Michael L.**; 2040 Main Street , 14th Floor, Irvine, California 92614 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, 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, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

(54) Title: CEA6 BINDING MOLECULES AND USES THEREOF

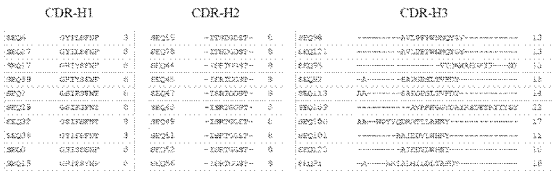


FIG. 1

(57) Abstract: Disclosed herein are carcinoembryonic antigen 6 (CEA6)-specific binding polypeptides. These binding polypeptides may be incorporated into chimeric antigen receptors (CARs). Also disclosed herein are methods of using these binding polypeptides and/or CARs for the treatment of, for example, a cancer.

[Continued on next page]

WO 2022/133089 A1

Declarations under Rule 4.17:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

Published:

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

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

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

CEA6 BINDING MOLECULES AND USES THEREOF

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit of priority of U.S. Provisional Patent Application No. 63/127885, filed December 18, 2020, and U.S. Provisional Patent Application No. 63/262312, filed October 8, 2021, each of which is hereby expressly incorporated by reference in its entirety.

REFERENCE TO SEQUENCE LISTING

[0002] The present application is being filed along with a Sequence Listing in electronic format. The Sequence Listing is provided as a file entitled BWGB005_SeqListing.TXT, which was created and last modified on December 16, 2021, which is 76,529 bytes in size. The information in the electronic Sequence Listing is hereby incorporated by reference in its entirety.

FIELD

[0003] Aspects of the present disclosure relate generally to CEA6 (CEACAM6)-specific binding polypeptides. These binding polypeptides can be incorporated into chimeric antigen receptor (CAR) constructs to be expressed in immune cells. These binding polypeptides and CARs may be used in the treatment of cancer.

BACKGROUND

[0004] Chimeric antigen receptor (CAR) T cells and other adoptive cell therapies have been shown to be effective in the treatment of cancer. The CAR, which is made up of an extracellular antigen binding domain, a transmembrane domain, and an intracellular signaling domain, enables directed killing of cancer cells based on cell surface antigen expression while minimally affecting normal cells that are not expressing the targeted antigen. The extracellular antigen binding domain is often made up of an antibody or a binding fragment or derivative thereof, such as a single chain variable fragment (scFv) or single domain antibody (sdAb). There is a present need for improved extracellular antigen binding domains to be used in CARs for the treatment of various cancers or other diseases.

SUMMARY OF THE DISCLOSURE

[0005] Disclosed herein are binding polypeptides that are able to bind to carcinoembryonic antigen 6 (CEA6; carcinoembryonic antigen-related cell adhesion molecule 6; CEACAM6). These binding polypeptides may be incorporated in a chimeric antigen receptor (CAR), which can be expressed by a cell. In some embodiments, the binding polypeptides are single domain antibodies (sdAbs).

[0006] Disclosed herein in some embodiments are CEA6 binding polypeptides comprising an immunoglobulin heavy chain variable domain comprising a CDR-H1, CDR-H2, and CDR-H3. In some embodiments, the CDR-H1 comprises a sequence having at least 90%, 95%, 99%, or 100% sequence identity to a sequence selected from **SEQ ID NOs: 1-43**. In some embodiments, CDR-H2 comprises a sequence having at least 90%, 95%, 99%, or 100% sequence identity to a sequence selected from **SEQ ID NOs: 44-86**. In some embodiments, the CDR-H3 comprises a sequence having at least 90%, 95%, 99%, or 100% sequence identity to a sequence selected from **SEQ ID NOs: 87-129**. In some embodiments, the CDR-H1 comprises a sequence having at least 90%, 95%, 99%, or 100% sequence identity to a sequence selected from **SEQ ID NOs: 1-43**, the CDR-H2 comprises a sequence having at least 90%, 95%, 99%, or 100% sequence identity to a sequence selected from **SEQ ID NOs: 44-86**, and the CDR-H3 comprises a sequence having at least 90%, 95%, 99%, or 100% sequence identity to a sequence selected from **SEQ ID NOs: 87-129**. In some embodiments, the immunoglobulin heavy chain variable domain comprises a sequence having at least 90%, 95%, 99%, or 100% sequence identity to a sequence selected from **SEQ ID NOs: 130-172**.

[0007] Also disclosed herein are nucleic acids that encode for any one of the CEA6 binding polypeptides disclosed herein.

[0008] Also disclosed herein are methods of treating a cancer in a subject in need thereof. In some embodiments, the methods comprise administering a chimeric antigen receptor cell to the subject. In some embodiments, the chimeric antigen receptor cell is any one of the chimeric antigen receptor cells disclosed herein. In some embodiments, the chimeric antigen receptor cell comprises any one or more of the CEA6 binding polypeptides disclosed herein.

[0009] Embodiments of the present invention provided herein are described by way of the following numbered alternatives:

[0010] 1. A carcinoembryonic antigen 6 (CEA6) binding polypeptide comprising an immunoglobulin heavy chain variable domain comprising a CDR-H1, CDR-H2, and CDR-H3, wherein:

the CDR-H1 comprises a sequence having at least 90%, 95%, 99%, or 100% sequence identity to a sequence selected from **SEQ ID NOs: 1-43**;

the CDR-H2 comprises a sequence having at least 90%, 95%, 99%, or 100% sequence identity to a sequence selected from **SEQ ID NOs: 44-86**; and

the CDR-H3 comprises a sequence having at least 90%, 95%, 99%, or 100% sequence identity to a sequence selected from **SEQ ID NOs: 87-129**.

[0011] 2. The CEA6 binding polypeptide of alternative 1, wherein:

1) the CDR-H1 comprises the sequence of **SEQ ID NO: 1**, the CDR-H2 comprises the sequence of **SEQ ID NO: 44**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 87**;

2) the CDR-H1 comprises the sequence of **SEQ ID NO: 2**, the CDR-H2 comprises the sequence of **SEQ ID NO: 45**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 88**;

3) the CDR-H1 comprises the sequence of **SEQ ID NO: 3**, the CDR-H2 comprises the sequence of **SEQ ID NO: 46**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 89**;

4) the CDR-H1 comprises the sequence of **SEQ ID NO: 4**, the CDR-H2 comprises the sequence of **SEQ ID NO: 47**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 90**;

5) the CDR-H1 comprises the sequence of **SEQ ID NO: 5**, the CDR-H2 comprises the sequence of **SEQ ID NO: 48**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 91**;

6) the CDR-H1 comprises the sequence of **SEQ ID NO: 6**, the CDR-H2 comprises the sequence of **SEQ ID NO: 49**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 92**;

7) the CDR-H1 comprises the sequence of **SEQ ID NO: 7**, the CDR-H2 comprises the sequence of **SEQ ID NO: 50**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 93**;

8) the CDR-H1 comprises the sequence of **SEQ ID NO: 8**, the CDR-H2 comprises the sequence of **SEQ ID NO: 51**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 94**;

9) the CDR-H1 comprises the sequence of **SEQ ID NO: 9**, the CDR-H2 comprises the sequence of **SEQ ID NO: 52**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 95**;

10) the CDR-H1 comprises the sequence of **SEQ ID NO: 10**, the CDR-H2 comprises the sequence of **SEQ ID NO: 53**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 96**;

11) the CDR-H1 comprises the sequence of **SEQ ID NO: 11**, the CDR-H2 comprises the sequence of **SEQ ID NO: 54**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 97**;

12) the CDR-H1 comprises the sequence of **SEQ ID NO: 12**, the CDR-H2 comprises the sequence of **SEQ ID NO: 55**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 98**;

13) the CDR-H1 comprises the sequence of **SEQ ID NO: 13**, the CDR-H2 comprises the sequence of **SEQ ID NO: 56**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 99**;

14) the CDR-H1 comprises the sequence of **SEQ ID NO: 14**, the CDR-H2 comprises the sequence of **SEQ ID NO: 57**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 100**;

15) the CDR-H1 comprises the sequence of **SEQ ID NO: 15**, the CDR-H2 comprises the sequence of **SEQ ID NO: 58**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 101**;

16) the CDR-H1 comprises the sequence of **SEQ ID NO: 16**, the CDR-H2 comprises the sequence of **SEQ ID NO: 59**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 102**;

17) the CDR-H1 comprises the sequence of **SEQ ID NO: 17**, the CDR-H2 comprises the sequence of **SEQ ID NO: 60**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 103**;

18) the CDR-H1 comprises the sequence of **SEQ ID NO: 18**, the CDR-H2 comprises the sequence of **SEQ ID NO: 61**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 104**;

19) the CDR-H1 comprises the sequence of **SEQ ID NO: 19**, the CDR-H2 comprises the sequence of **SEQ ID NO: 62**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 105**;

20) the CDR-H1 comprises the sequence of **SEQ ID NO: 20**, the CDR-H2 comprises the sequence of **SEQ ID NO: 63**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 106**;

21) the CDR-H1 comprises the sequence of **SEQ ID NO: 21**, the CDR-H2 comprises the sequence of **SEQ ID NO: 64**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 107**;

22) the CDR-H1 comprises the sequence of **SEQ ID NO: 22**, the CDR-H2 comprises the sequence of **SEQ ID NO: 65**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 108**;

23) the CDR-H1 comprises the sequence of **SEQ ID NO: 23**, the CDR-H2 comprises the sequence of **SEQ ID NO: 66**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 109**;

24) the CDR-H1 comprises the sequence of **SEQ ID NO: 24**, the CDR-H2 comprises the sequence of **SEQ ID NO: 67**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 110**;

25) the CDR-H1 comprises the sequence of **SEQ ID NO: 25**, the CDR-H2 comprises the sequence of **SEQ ID NO: 68**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 111;**

26) the CDR-H1 comprises the sequence of **SEQ ID NO: 26**, the CDR-H2 comprises the sequence of **SEQ ID NO: 69**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 112**;

27) the CDR-H1 comprises the sequence of **SEQ ID NO: 27**, the CDR-H2 comprises the sequence of **SEQ ID NO: 70**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 113**;

28) the CDR-H1 comprises the sequence of **SEQ ID NO: 28**, the CDR-H2 comprises the sequence of **SEQ ID NO: 71**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 114**;

29) the CDR-H1 comprises the sequence of **SEQ ID NO: 29**, the CDR-H2 comprises the sequence of **SEQ ID NO: 72**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 115**;

30) the CDR-H1 comprises the sequence of **SEQ ID NO: 30**, the CDR-H2 comprises the sequence of **SEQ ID NO: 73**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 116**;

31) the CDR-H1 comprises the sequence of **SEQ ID NO: 31**, the CDR-H2 comprises the sequence of **SEQ ID NO: 74**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 117**;

32) the CDR-H1 comprises the sequence of **SEQ ID NO: 32**, the CDR-H2 comprises the sequence of **SEQ ID NO: 75**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 118**;

33) the CDR-H1 comprises the sequence of **SEQ ID NO: 33**, the CDR-H2 comprises the sequence of **SEQ ID NO: 76**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 119**;

34) the CDR-H1 comprises the sequence of **SEQ ID NO: 34**, the CDR-H2 comprises the sequence of **SEQ ID NO: 77**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 120**;

35) the CDR-H1 comprises the sequence of **SEQ ID NO: 35**, the CDR-H2 comprises the sequence of **SEQ ID NO: 78**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 121**;

36) the CDR-H1 comprises the sequence of **SEQ ID NO: 36**, the CDR-H2 comprises the sequence of **SEQ ID NO: 79**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 122**;

37) the CDR-H1 comprises the sequence of **SEQ ID NO: 37**, the CDR-H2 comprises the sequence of **SEQ ID NO: 80**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 123**;

38) the CDR-H1 comprises the sequence of **SEQ ID NO: 38**, the CDR-H2 comprises the sequence of **SEQ ID NO: 81**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 124**;

39) the CDR-H1 comprises the sequence of **SEQ ID NO: 39**, the CDR-H2 comprises the sequence of **SEQ ID NO: 82**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 125**;

40) the CDR-H1 comprises the sequence of **SEQ ID NO: 40**, the CDR-H2 comprises the sequence of **SEQ ID NO: 83**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 126**;

41) the CDR-H1 comprises the sequence of **SEQ ID NO: 41**, the CDR-H2 comprises the sequence of **SEQ ID NO: 84**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 127**;

42) the CDR-H1 comprises the sequence of **SEQ ID NO: 42**, the CDR-H2 comprises the sequence of **SEQ ID NO: 85**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 128**; or

43) the CDR-H1 comprises the sequence of **SEQ ID NO: 43**, the CDR-H2 comprises the sequence of **SEQ ID NO: 86**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 129**.

[0012] 3. The CEA6 binding polypeptide of alternative 1 or 2, wherein the heavy chain variable domain comprises an amino acid sequence having at least 90%, 95%, 99%, or 100% sequence identity to any sequence selected from **SEQ ID NOs: 130-172**.

[0013] 4. The CEA6 binding polypeptide of any one of alternatives 1-3, wherein the CEA6 binding polypeptide is humanized.

[0014] 5. The CEA6 binding polypeptide of any one of alternatives 1-4, wherein the CEA6 binding polypeptide is a single domain antibody (sdAb).

[0015] 6. A chimeric antigen receptor (CAR) comprising the CEA6 binding polypeptide of any one of alternatives 1-5.

[0016] 7. A chimeric antigen receptor (CAR) cell comprising the CAR of alternative 6.

[0017] 8. The CAR cell of alternative 7, wherein the CAR cell is a CAR T cell.

[0018] 9. The CAR cell of alternative 7 or 8, wherein the CAR cell comprises at least two binding polypeptides and the CAR cell is a multivalent CAR cell.

[0019] 10. The CAR cell of any one of alternatives 7-9, wherein the CAR cell is derived from a subject or from a cell line.

[0020] 11. The CAR cell of alternative 10, wherein the subject has a cancer.

[0021] 12. The CAR cell of alternative 11, wherein the cancer is breast cancer, colorectal cancer, kidney cancer, liver cancer, lung cancer, brain cancer, pancreatic cancer, bladder cancer, testicular cancer, prostate cancer, gastric cancer, ovarian cancer, head and neck cancer, gallbladder cancer, a hematologic malignancy, or any combination thereof.

[0022] 13. A nucleic acid that encodes for a polypeptide comprising a sequence having at least 90%, 95%, 99%, or 100% sequence identity to the CEA6 binding polypeptide of any one of alternatives 1-5 or the CAR of alternative 6.

[0023] 14. A method of treating a cancer in a subject in need thereof, comprising administering the CAR cell of any one of alternatives 7-12.

[0024] 15. The method of alternative 14, wherein the chimeric antigen receptor cell is autologous or allogeneic to the subject.

[0025] 16. The method of alternative 14 or 15, wherein the subject is a mammal.

[0026] 17. The method of any one of alternatives 14-16, wherein the subject is a human.

[0027] 18. The method of any one of alternatives 14-17, wherein the cancer is breast cancer, colorectal cancer, kidney cancer, liver cancer, lung cancer, brain cancer, pancreatic cancer, bladder cancer, testicular cancer, prostate cancer, gastric cancer, ovarian cancer, head and neck cancer, gallbladder cancer, a hematologic malignancy, or any combination thereof.

[0028] 19. The method of any one of alternatives 14-18, wherein the chimeric antigen receptor cell is administered parenterally.

BRIEF DESCRIPTION OF THE DRAWINGS

[0029] In addition to the features described above, additional features and variations will be readily apparent from the following descriptions of the drawings and exemplary embodiments. It is to be understood that these drawings depict typical embodiments and are not intended to be limiting in scope.

[0030] **FIG. 1** depicts an exemplary alignment for the heavy chain variable domain CDRs disclosed herein.

[0031] **FIGs. 2A and 2B** are bar charts which show the results of cytotoxicity assays of exemplary anti-CEA6 CAR T-cell lines against MKN-45 gastric cancer cells, where the T-cells have been obtained from two donors, Z0016 (**FIG. 2A**) or Y1287 (**FIG. 2B**).

[0032] **FIGs. 2C and 2D** are bar charts which show the results of cytotoxicity assays of exemplary anti-CEA6 CAR T-cell lines against BxPC-3 pancreatic cancer cells, where the T-cells have been obtained from two donors Z0016 (**FIG. 2C**) or Y1287 (**FIG. 2D**).

[0033] **FIGs. 2E and 2F** are bar charts which show the results of cytotoxicity assays of exemplary anti-CEA6 CAR T-cell lines against Capan-1 pancreatic cancer cells, where the T-cells have been obtained from two donors Z0016 (**FIG. 2E**) or Y1287 (**FIG. 2F**).

[0034] **FIG. 3A** shows three line graphs which show the anti-tumor efficacy of an exemplary anti-CEA6 CAR T-cell line B4T2-001 ('001) against tumors in a cell line-derived xenograft (CDX) model of gastric tumor , non-small cell lung cancer (NSCLC), and pancreatic tumor.

[0035] **FIG. 3B** show an image of a pancreatic orthotopic model of pancreatic tumors and also shows a line graph detailing the efficacy of an exemplary anti-CEA6 CAR T-cell line B4T2-001 ('001) against tumors in the pancreatic orthotopic and patient-derived colorectal cancer xenograph model.

[0036] **FIG. 4** is a line graph showing the results of a tumor re-challenge study showing immunological memory of an exemplary anti-CEA6 CAR T-cell line B4T2-001 that drives a secondary complete response.

DETAILED DESCRIPTION OF THE DISCLOSURE

[0037] Disclosed herein are binding polypeptides that are incorporated into a chimeric antigen receptor cell. In some embodiments, the chimeric antigen receptor cell is a chimeric antigen receptor T cell (CAR-T cell). These CAR-Ts may be constructed through processes conventionally known in the art. The binding polypeptides provide specificity towards their respective tumor-associated antigens, enabling targeting of cancers expressing said tumor-associated antigens by the CAR-T cell.

[0038] In some embodiments, the binding polypeptides are single domain antibodies (sdAbs) disposed on the surface of the chimeric antigen receptor cells (e.g. CAR-T cell). The sdAbs may be specific for, or have binding affinity towards, a tumor-associated antigen. In some embodiments, the tumor-associated antigen is carcinoembryonic antigen 6 (CEA6).

[0039] Also disclosed herein are methods of treating a cancer in a subject in need thereof by administering a chimeric antigen receptor cell comprising one or more of the binding polypeptides disclosed herein. In some embodiments, the cancer may be breast cancer, colorectal cancer, kidney cancer, liver cancer, lung cancer, brain cancer, pancreatic cancer, bladder cancer, testicular cancer, prostate cancer, gastric cancer, ovarian cancer, head and neck cancer, gallbladder cancer, a hematologic malignancy, or any combination thereof. In some embodiments, the

hematologic malignancy may comprise leukemia, acute lymphoblastic leukemia, acute myelogenous leukemia, chronic myelogenous leukemia, chronic lymphocytic leukemia, hairy cell leukemia, lymphoma, Hodgkin's disease, Non-Hodgkin lymphoma, or multiple myeloma. The CAR-T cell may be derived from the subject for an autologous treatment. Alternatively, the CAR-T cell may be derived from the same species as the subject for an allogeneic treatment.

Definitions

[0040] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as is commonly understood by one of skill in the art to which the claimed subject matter belongs. It is to be understood that the foregoing general description and the following detailed description are exemplary and explanatory only and are not restrictive of any subject matter claimed.

[0041] The section headings used herein are for organizational purposes only and are not to be construed as limiting the subject matter described.

[0042] The articles "a" and "an" are used herein to refer to one or to more than one (for example, at least one) of the grammatical object of the article. By way of example, "an element" means one element or more than one element.

[0043] By "about" is meant a quantity, level, value, number, frequency, percentage, dimension, size, amount, weight or length that varies by as much as 30, 25, 20, 15, 10, 9, 8, 7, 6, 5, 4, 3, 2 or 1% to a reference quantity, level, value, number, frequency, percentage, dimension, size, amount, weight or length.

[0044] Throughout this specification, unless the context requires otherwise, the words "comprise," "comprises," and "comprising" will be understood to imply the inclusion of a stated step or element or group of steps or elements but not the exclusion of any other step or element or group of steps or elements. By "consisting of" is meant including, and limited to, whatever follows the phrase "consisting of." Thus, the phrase "consisting of" indicates that the listed elements are required or mandatory, and that no other elements may be present. By "consisting essentially of" is meant including any elements listed after the phrase and limited to other elements that do not interfere with or contribute to the activity or action specified in the disclosure for the listed elements. Thus, the phrase "consisting essentially of" indicates that the listed elements are required or mandatory, but that other elements are optional and may or may not be present depending upon whether or not they materially affect the activity or action of the listed elements.

[0045] As used herein, the terms "individual(s)", "subject(s)" and "patient(s)" mean any mammal. In some embodiments, the mammal is a human. In some embodiments, the mammal is a non-human. None of the terms require or are limited to situations characterized by the

supervision (e.g. constant or intermittent) of a health care worker (e.g. a doctor, a registered nurse, a nurse practitioner, a physician's assistant, an orderly or a hospice worker).

[0046] The term "administering" includes enteral, oral, intranasal, parenteral, intravenous, intraperitoneal, intramuscular, intra-arteriole, intraventricular, intradermal, intralesional, intracranial, intrathecal, or subcutaneous administration, or the implantation of a slow-release device, e.g., a mini-osmotic pump, to a subject.

[0047] The terms "nucleic acid" or "nucleic acid molecule" as used herein refers to polynucleotides, such as deoxyribonucleic acid (DNA) or ribonucleic acid (RNA), oligonucleotides, fragments generated by the polymerase chain reaction (PCR), and fragments generated by any of ligation, scission, endonuclease action, and exonuclease action. Nucleic acid molecules can be composed of monomers that are naturally-occurring nucleotides (such as DNA and RNA), or analogs of naturally-occurring nucleotides (e.g., enantiomeric forms of naturally-occurring nucleotides), or a combination of both. Nucleic acid monomers can be linked by phosphodiester bonds or analogs of such linkages. A nucleic acid or nucleic acids can be contained in a nucleic acid vector or nucleic acid construct (e.g. plasmid, virus, bacteriophage, cosmid, fosmid, phagemid, bacterial artificial chromosome (BAC), yeast artificial chromosome (YAC), or human artificial chromosome (HAC)) that can be used for amplification and/or expression of the nucleic acid or nucleic acids in various biological systems. Typically, the vector or construct will also contain elements including but not limited to promoters, enhancers, terminators, inducers, ribosome binding sites, translation initiation sites, start codons, stop codons, polyadenylation signals, origins of replication, cloning sites, multiple cloning sites, restriction enzyme sites, epitopes, reporter genes, selection markers, antibiotic selection markers, targeting sequences, peptide purification tags, or accessory genes, or any combination thereof.

[0048] A nucleic acid or nucleic acid molecule can comprise one or more sequences encoding different peptides, polypeptides, or proteins. These one or more sequences can be joined in the same nucleic acid or nucleic acid molecule adjacently, or with extra nucleic acids in between, e.g. linkers, repeats or restriction enzyme sites, or any other sequence that is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, or 300 bases long, or any length in a range defined by any two of the aforementioned lengths. The term "downstream" on a nucleic acid as used herein refers to a sequence being after the 3'-end of a previous sequence, on the strand containing the encoding sequence (sense strand) if the nucleic acid is double stranded. The term "upstream" on a nucleic acid as used herein refers to a sequence being before the 5'-end of a subsequent sequence, on the strand containing the encoding sequence (sense strand) if the nucleic acid is double stranded. The term "grouped" on a nucleic acid as used herein refers to two or more sequences that occur in

proximity either directly or with extra nucleic acids in between, e.g. linkers, repeats, or restriction enzyme sites, or any other sequence that is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, or 300 bases long, or any length in a range defined by any two of the aforementioned lengths, but generally not with a sequence in between that encodes for a functioning or catalytic polypeptide, protein, or protein domain.

[0049] The term "codon optimized" regarding a nucleic acid as used herein refers to the substitution of codons of the nucleic acid to enhance or maximize translation in a host of a particular species without changing the polypeptide sequence based on species-specific codon usage biases and relative availability of each aminoacyl-tRNA in the target cell cytoplasm. Codon optimization and techniques to perform such optimization is known in the art. Those skilled in the art will appreciate that gene expression levels are dependent on many factors, such as promoter sequences and regulatory elements. In this aspect, many synthetic genes can be designed to increase their protein expression level.

[0050] The terms "peptide", "polypeptide", and "protein" as used herein refers to macromolecules comprised of amino acids linked by peptide bonds. The numerous functions of peptides, polypeptides, and proteins are known in the art, and include but are not limited to enzymes, structure, transport, defense, hormones, or signaling. Peptides, polypeptides, and proteins are often, but not always, produced biologically by a ribosomal complex using a nucleic acid template, although chemical syntheses are also available. By manipulating the nucleic acid template, peptide, polypeptide, and protein mutations such as substitutions, deletions, truncations, additions, duplications, or fusions of more than one peptide, polypeptide, or protein can be performed. These fusions of more than one peptide, polypeptide, or protein can be joined in the same molecule adjacently, or with extra amino acids in between, e.g. linkers, repeats, epitopes, or tags, or any other sequence that is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, or 300 bases long, or any length in a range defined by any two of the aforementioned lengths. The term "downstream" on a polypeptide as used herein refers to a sequence being after the C-terminus of a previous sequence. The term "upstream" on a polypeptide as used herein refers to a sequence being before the N-terminus of a subsequent sequence.

[0051] In some embodiments, the nucleic acid or peptide sequences presented herein and used in the examples are functional in various biological systems including but not limited to humans, mice, rats, monkeys, primates, cats, dogs, rabbits, *E. coli*, yeast, and mammalian cells. In other embodiments, nucleic acid or peptide sequences sharing at least or lower than 0%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%,

or 100% identity, or any percentage within a range defined by any two of the aforementioned percentages of identity to the nucleic acid or peptide sequences presented herein and used in the examples can also be used with little or no effect on the function of the sequences in biological systems. As used herein, the term "identity" refers to a nucleic acid or peptide sequence having the same overall order of nucleotide or amino acids, respectively, as a template nucleic acid or peptide sequence with specific changes such as substitutions, deletions, repetitions, or insertions within the sequence. In some embodiments, two nucleic acid sequences sharing as low as 0%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity can encode for the same polypeptide by comprising different codons that encode for the same amino acid during translation.

[0052] As disclosed herein, sequences having a % homology to any of the sequences disclosed herein are envisioned and may be used. The term "% homology" refers to the degree of conservation between two sequences when considering their three-dimensional structure. For example, homology between two protein sequences may be dependent on structural motifs, such as beta strands, alpha helices, and other folds, as well as their distribution throughout the sequence. Homology may be determined through structural determination, either empirically or *in silico*. In some embodiments, any sequence having at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence homology to any of the sequences disclosed herein may be used. In some embodiments, any sequence having at least 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 substitutions, deletions, or additions relative to any of the sequences disclosed herein, which may or may not affect the overall % homology, may be used.

[0053] As applied herein, sequences having a certain % similarity to any of the sequence disclosed herein are envisioned and may be used. In some embodiments, these sequences may include peptide sequences, nucleic acid sequences, CDR sequences, variable region sequences, or heavy or light chain sequences. As understood in the art with respect to peptide sequences, "similarity" refers to the comparison of amino acids based on their properties, including but not limited to size, polarity, charge, pK, aromaticity, hydrogen bonding properties, or presence of functional groups (e.g. hydroxyl, thiol, amine, carboxyl, and the like). The term "% similarity" refers to the percentage of units (i.e. amino acids) that are the same between two or more sequences relative to the length of the sequence. When the two or more sequences being compared are the same length, the % similarity will be respective that length. When two or more sequences being compared are different lengths, deletions and/or insertions may be introduced to obtain the best alignment. The similarity of two amino acids may dictate whether a certain

substitution is conservative or non-conservative. Methods of determining the conservativeness of an amino acid substitution are generally known in the art and may involve substitution matrices. Commonly used substitution matrices include BLOSUM45, BLOSUM62, BLOSUM80, PAM100, PAM120, PAM160, PAM200, PAM250, but other substitution matrices or approaches may be used as considered appropriate by the skilled person. A certain substitution matrix may be preferential over the others when considering aspects such as stringency, conservation and/or divergence of related sequences (e.g. within the same species or broader), and length of the sequences in question. As used herein, a peptide sequence having a certain % similarity to another sequence will have up to that % of amino acids that are either identical or an acceptable substitution as governed by the method of similarity determination used. In some embodiments, a sequence having at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence similarity to any of the sequences disclosed herein may be used. In some embodiments, any sequence having at least 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 similar substitutions relative to any of the sequences disclosed herein may be used. As applied to antibody sequences, these similar substitutions may apply to antigen-binding regions (i.e. CDRs) or regions that do not bind to antigens or are only secondary to antigen binding (i.e. framework regions).

[0054] The term “consensus sequence” as used herein with regard to sequences refers to the generalized sequence representing all of the different combinations of permissible amino acids at each location of a group of sequences. A consensus sequence may provide insight into the conserved regions of related sequences where the unit (e.g. amino acid or nucleotide) is the same in most or all of the sequences, and regions that exhibit divergence between sequences. In the case of antibodies, the consensus sequence of a CDR may indicate amino acids that are important or dispensable for antigen binding. It is envisioned that consensus sequences may be prepared with any of the sequences provided herein, and the resultant various sequences derived from the consensus sequence can be validated to have similar effects as the template sequences.

[0055] As used herein, the term "antibody" denotes the meaning ascribed to it by one of skill in the art, and further it is intended to include any polypeptide chain-containing molecular structure with a specific shape that fits to and recognizes an epitope, where one or more non-covalent binding interactions stabilize the complex between the molecular structure and the epitope.

[0056] The term "antibody library" refers to a collection of antibodies and/or antibody fragments displayed for screening and/or combination into full antibodies. The antibodies and/or

antibody fragments may be displayed on a ribosome; on a phage; or on a cell surface, in particular a yeast cell surface.

[0057] The term "compete," as used herein with regard to an antibody or binding polypeptide, means that a first antibody or binding polypeptide, or an antigen-binding portion thereof, binds to an epitope in a manner sufficiently similar to the binding of a second antibody or binding polypeptide, or an antigen-binding portion thereof, such that the result of binding of the first antibody or binding polypeptide with its cognate epitope is detectably decreased in the presence of the second antibody or binding polypeptide compared to the binding of the first antibody or binding polypeptide in the absence of the second antibody or binding polypeptide. The alternative, where the binding of the second antibody or binding polypeptide to its epitope is also detectably decreased in the presence of the first antibody or binding polypeptide, can, but need not be the case. Regardless of the mechanism by which such competition occurs (e.g., steric hindrance, conformational change, or binding to a common epitope, or portion thereof), the skilled artisan would appreciate, based upon the teachings provided herein, that such competing antibodies or binding polypeptides are encompassed and can be useful for the methods disclosed herein.

[0058] An antibody or binding polypeptide that "preferentially binds" or "specifically binds" (used interchangeably herein) to an epitope is a term well understood in the art, and methods to determine such specific or preferential binding are also well known in the art. A molecule is said to exhibit "specific binding" or "preferential binding" if it reacts or associates more frequently, and/or more rapidly, and/or with greater duration and/or with greater affinity with a particular cell or substance than it does with alternative cells or substances. An antibody or binding polypeptide "specifically binds" or "preferentially binds" to a target if it binds with greater affinity, and/or avidity, and/or more readily, and/or with greater duration than it binds to other substances.

[0059] The term “humanized” as applies to a non-human (e.g. rodent or primate) antibodies are hybrid immunoglobulins, immunoglobulin chains or fragments thereof which contain minimal sequence derived from non-human immunoglobulin.

[0060] The term “single domain binding polypeptide” or “single domain antibody” (sdAb) as used herein refers to a single peptide strand (e.g. not bound to another peptide strand with disulfide bonds) comprising an intact immunoglobulin domain or other protein fold which can recognize antigens. Single domain binding polypeptides or sdAbs may be derived from typical heavy or light immunoglobulin chains, such as from human, or from alternative sources such as dromedaries (e.g. V_HH) and cartilaginous fish (e.g. V_{NAR}). In some embodiments, the single domain binding polypeptide or sdAb comprises one, two, or three complementarity determining

regions (CDRs). In some embodiments, the single domain binding polypeptide or sdAb comprises one, two, or three of a CDR1, CDR2, and CDR3.

[0061] The term "single-chain variable fragment" (scFv) as used herein is a fusion protein comprising the variable regions of the heavy (VH) and light chains (VL) of an immunoglobulin, in which the VH and VL are covalently linked to form a VH:VL heterodimer. The VH and VL are either joined directly or joined by a peptide-encoding linker, which connects the N-terminus of the VH with the C-terminus of the VL, or the C-terminus of the VH with the N-terminus of the VL. The linker is usually rich in glycine for flexibility, as well as serine or threonine for solubility. Despite removal of the constant regions and the introduction of a linker, scFv proteins retain the specificity of the original immunoglobulin. Single chain Fv polypeptide antibodies can be expressed from a nucleic acid including VH- and VL-encoding sequences. In some embodiments, the VH and VL of the scFv each comprises one, two, or three CDRs. In some embodiments, the VH and VL of the scFv each comprises one, two, or three of a CDR1, CDR2, and CDR3.

[0062] In certain embodiments, definitive delineation of a CDR and identification of residues comprising the binding site of an antibody or binding polypeptide is accomplished by solving the structure of the antibody or binding polypeptide and/or solving the structure of the antibody-ligand complex. In certain embodiments, that can be accomplished by any of a variety of techniques known to those skilled in the art, such as X-ray crystallography. In certain embodiments, various methods of analysis can be employed to identify or approximate the CDR regions. In certain embodiments, various methods of analysis can be employed to identify or approximate the CDR regions. Examples of such methods include, but are not limited to, the Kabat definition, the Chothia definition, the IMGT approach (Lefranc et al., 2003) Dev Comp Immunol. 27:55-77), computational programs such as Paratome (Kunik et al., 2012, Nucl Acids Res. W521-4), the AbM definition, and the conformational definition.

[0063] The Kabat definition is a standard for numbering the residues in an antibody and is typically used to identify CDR regions. See, e.g., Johnson & Wu, 2000, Nucleic Acids Res., 28: 214-8. The Chothia definition is similar to the Kabat definition, but the Chothia definition takes into account positions of certain structural loop regions. See, e.g., Chothia et al., 1986, J. Mol. Biol., 196: 901-17; Chothia et al., 1989, Nature, 342: 877-83. The AbM definition uses an integrated suite of computer programs produced by Oxford Molecular Group that model antibody structure. See, e.g., Martin et al., 1989, Proc Natl Acad Sci (USA), 86:9268-9272; "AbM.TM., A Computer Program for Modeling Variable Regions of Antibodies," Oxford, UK; Oxford Molecular, Ltd. The AbM definition models the tertiary structure of an antibody from primary sequence using a combination of knowledge databases and ab initio methods, such as those

described by Samudrala et al., 1999, "Ab Initio Protein Structure Prediction Using a Combined Hierarchical Approach," in PROTEINS, Structure, Function and Genetics Suppl., 3:194-198. The contact definition is based on an analysis of the available complex crystal structures. See, e.g., MacCallum et al., 1996, J. Mol. Biol., 5:732-45. In another approach, referred to herein as the "conformational definition" of CDRs, the positions of the CDRs may be identified as the residues that make enthalpic contributions to antigen binding. See, e.g., Makabe et al., 2008, Journal of Biological Chemistry, 283:1156-1166. Still other CDR boundary definitions may not strictly follow one of the above approaches, but will nonetheless overlap with at least a portion of the Kabat CDRs, although they may be shortened or lengthened in light of prediction or experimental findings that particular residues or groups of residues do not significantly impact antigen binding. As used herein, a CDR may refer to CDRs defined by any approach known in the art, including combinations of approaches. The methods used herein may utilize CDRs defined according to any of these approaches. For any given embodiment containing more than one CDR, the CDRs may be defined in accordance with any of Kabat, Chothia, extended, IMGT, Paratome, AbM, and/or conformational definitions, or a combination of any of the foregoing.

[0064] The term “chimeric antigen receptor (CAR)” as used herein refers to engineered biological receptors that confers an artificial specificity in an immune cell towards a certain antigen, such as a tumor-associated antigen. An exemplary immune cell in which CARs can be used are T cells, but it is envisioned that CARs can be engineered into any amenable cytotoxic immune cell, including but not limited to T cells, Natural Killer (NK) cells, Natural Killer T (NKT) cells, dendritic cells, or macrophages. In this aspect, any disclosure pertaining to CAR T cells can also be applied to other immune cells comprising CARs. At their core, CARs comprise an extracellular antigen-recognizing domain (e.g. tumor receptor ligand, or antibody), hinge, transmembrane, and intracellular signaling domain (endodomain). Different combinations of these CAR components may result in different specificities and efficacy against certain cancer antigens.

[0065] As used herein, the terms “treating” or “treatment” (and as well understood in the art) means an approach for obtaining beneficial or desired results in a subject's condition, including clinical results. Beneficial or desired clinical results can include, but are not limited to, alleviation or amelioration of one or more symptoms or conditions, diminishment of the extent of a disease, stabilizing (i.e., not worsening) the state of disease, prevention of a disease's transmission or spread, delaying or slowing of disease progression, amelioration or palliation of the disease state, diminishment of the reoccurrence of disease, and remission, whether partial or total and whether detectable or undetectable. “Treating” and “treatment” as used herein also include prophylactic treatment. Treatment methods comprise administering to a subject a

therapeutically effective amount of an active agent. The administering step may consist of a single administration or may comprise a series of administrations. The compositions are administered to the subject in an amount and for a duration sufficient to treat the subject. The length of the treatment period depends on a variety of factors, such as the severity of the condition, the age and genetic profile of the subject, the concentration of active agent, the activity of the compositions used in the treatment, or a combination thereof. It will also be appreciated that the effective dosage of an agent used for the treatment or prophylaxis may increase or decrease over the course of a particular treatment or prophylaxis regime. Changes in dosage may result and become apparent by standard diagnostic assays known in the art. In some instances, chronic administration may be required.

[0066] The terms “effective amount” or “effective dose” as used herein refers to that amount of a recited composition or compound that results in an observable designated effect. Actual dosage levels of active ingredients in an active composition of the presently disclosed subject matter can be varied so as to administer an amount of the active composition or compound that is effective to achieve the designated response for a particular subject and/or application. The selected dosage level can vary based upon a variety of factors including, but not limited to, the activity of the composition, formulation, route of administration, combination with other drugs or treatments, severity of the condition being treated, and the physical condition and prior medical history of the subject being treated. In some embodiments, a minimal dose is administered, and dose is escalated in the absence of dose-limiting toxicity to a minimally effective amount. Determination and adjustment of an effective dose, as well as evaluation of when and how to make such adjustments, are contemplated herein.

[0067] The term “administering” includes oral administration, topical contact, administration as a suppository, parenteral, intravenous, intraperitoneal, intramuscular, intralesional, intrathecal, intranasal, subdermal, or subcutaneous administration, or the implantation of a slow-release device, e.g., a mini-osmotic pump, to a subject. Administration is by any route, including parenteral and transmucosal (e.g., buccal, sublingual, palatal, gingival, nasal, vaginal, rectal, or transdermal). Parenteral administration includes, e.g., intravenous, intramuscular, intra-arteriole, intradermal, subcutaneous, intraperitoneal, intraventricular, and intracranial. Other modes of delivery include, but are not limited to, the use of liposomal formulations, intravenous infusion, transdermal patches, etc. By “co-administer” it is meant that a first compound described herein is administered at the same time, just prior to, or just after the administration of a second compound described herein.

[0068] As used herein, the term "therapeutic target" refers to a gene or gene product that, upon modulation of its activity (e.g., by modulation of expression, biological activity, and the

like), can provide for modulation of the disease phenotype. As used throughout, "modulation" is meant to refer to an increase or a decrease in the indicated phenomenon (e.g., modulation of a biological activity refers to an increase in a biological activity or a decrease in a biological activity).

[0069] As used herein, the term “standard of care”, “best practice” and “standard therapy” refers to the treatment that is accepted by medical practitioners to be an appropriate, proper, effective, and/or widely used treatment for a certain disease. The standard of care of a certain disease depends on many different factors, including the biological effect of treatment, region or location within the body, patient status (e.g. age, weight, gender, hereditary risks, other disabilities, secondary conditions), toxicity, metabolism, bioaccumulation, therapeutic index, dosage, and other factors known in the art. Determining a standard of care for a disease is also dependent on establishing safety and efficacy in clinical trials as standardized by regulatory bodies such as the US Food and Drug Administration, International Council for Harmonisation, Health Canada, European Medicines Agency, Therapeutics Goods Administration, Central Drugs Standard Control Organization, National Medical Products Administration, Pharmaceuticals and Medical Devices Agency, Ministry of Food and Drug Safety, and the World Health Organization. The standard of care for a disease may include but is not limited to surgery, radiation, chemotherapy, targeted therapy, or immunotherapy.

[0070] The term “% w/w” or “% wt/wt” means a percentage expressed in terms of the weight of the ingredient or agent over the total weight of the composition multiplied by 100.

Antigen binding polypeptides

[0071] Unless otherwise specified, the complementarity determining regions (CDRs) disclosed herein follow the IMGT definition. However, the CDRs, either separately or within the context of the variable domains, can also be interpreted by Kabat, Chothia, or other definitions as understood by those of skill in the art.

[0072] Disclosed herein are carcinoembryonic antigen 6 (carcinoembryonic antigen-related cell adhesion molecule 6; CEA6; CEACAM6) binding polypeptides. In some embodiments, the CEA6 binding polypeptides comprise an immunoglobulin heavy chain variable domain comprising a CDR-H1, CDR-H2, and CDR-H3. In some embodiments, the CDR-H1 comprises a sequence having at least 90%, 95%, 99%, or 100% sequence identity to a sequence selected from **SEQ ID NOs: 1-43**. In some embodiments, CDR-H2 comprises a sequence having at least 90%, 95%, 99%, or 100% sequence identity to a sequence selected from **SEQ ID NOs: 44-86**. In some embodiments, the CDR-H3 comprises a sequence having at least 90%, 95%, 99%, or 100% sequence identity to a sequence selected from **SEQ ID NOs: 87-129**. In some

embodiments, the CDR-H1 comprises a sequence having at least 90%, 95%, 99%, or 100% sequence identity to a sequence selected from **SEQ ID NOs: 1-43**, the CDR-H2 comprises a sequence having at least 90%, 95%, 99%, or 100% sequence identity to a sequence selected from **SEQ ID NOs: 44-86**, and the CDR-H3 comprises a sequence having at least 90%, 95%, 99%, or 100% sequence identity to a sequence selected from **SEQ ID NOs: 87-129**.

[0073] In some embodiments of the CEA6 binding polypeptides: 1) the CDR-H1 comprises the sequence of **SEQ ID NO: 1**, the CDR-H2 comprises the sequence of **SEQ ID NO: 44**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 87**; 2) the CDR-H1 comprises the sequence of **SEQ ID NO: 2**, the CDR-H2 comprises the sequence of **SEQ ID NO: 45**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 88**; 3) the CDR-H1 comprises the sequence of **SEQ ID NO: 3**, the CDR-H2 comprises the sequence of **SEQ ID NO: 46**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 89**; 4) the CDR-H1 comprises the sequence of **SEQ ID NO: 4**, the CDR-H2 comprises the sequence of **SEQ ID NO: 47**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 90**; 5) the CDR-H1 comprises the sequence of **SEQ ID NO: 5**, the CDR-H2 comprises the sequence of **SEQ ID NO: 48**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 91**; 6) the CDR-H1 comprises the sequence of **SEQ ID NO: 6**, the CDR-H2 comprises the sequence of **SEQ ID NO: 49**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 92**; 7) the CDR-H1 comprises the sequence of **SEQ ID NO: 7**, the CDR-H2 comprises the sequence of **SEQ ID NO: 50**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 93**; 8) the CDR-H1 comprises the sequence of **SEQ ID NO: 8**, the CDR-H2 comprises the sequence of **SEQ ID NO: 51**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 94**; 9) the CDR-H1 comprises the sequence of **SEQ ID NO: 9**, the CDR-H2 comprises the sequence of **SEQ ID NO: 52**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 95**; 10) the CDR-H1 comprises the sequence of **SEQ ID NO: 10**, the CDR-H2 comprises the sequence of **SEQ ID NO: 53**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 96**; 11) the CDR-H1 comprises the sequence of **SEQ ID NO: 11**, the CDR-H2 comprises the sequence of **SEQ ID NO: 54**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 7**; 12) the CDR-H1 comprises the sequence of **SEQ ID NO: 12**, the CDR-H2 comprises the sequence of **SEQ ID NO: 55**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 98**; 13) the CDR-H1 comprises the sequence of **SEQ ID NO: 13**, the CDR-H2 comprises the sequence of **SEQ ID NO: 56**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 99**; 14) the CDR-H1 comprises the sequence of **SEQ ID NO: 14**, the CDR-H2 comprises the sequence of **SEQ ID NO: 57**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 100**; 15) the CDR-H1 comprises the sequence of **SEQ ID NO: 15**, the CDR-H2 comprises the sequence of **SEQ ID NO: 58**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 101**; 16) the CDR-H1 comprises the sequence of **SEQ ID NO: 16**, the CDR-H2 comprises

the sequence of **SEQ ID NO: 59**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 102**; 17) the CDR-H1 comprises the sequence of **SEQ ID NO: 17**, the CDR-H2 comprises the sequence of **SEQ ID NO: 60**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 103**; 18) the CDR-H1 comprises the sequence of **SEQ ID NO: 18**, the CDR-H2 comprises the sequence of **SEQ ID NO: 61**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 104**; 19) the CDR-H1 comprises the sequence of **SEQ ID NO: 19**, the CDR-H2 comprises the sequence of **SEQ ID NO: 62**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 105**; 20) the CDR-H1 comprises the sequence of **SEQ ID NO: 20**, the CDR-H2 comprises the sequence of **SEQ ID NO: 63**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 106**; 21) the CDR-H1 comprises the sequence of **SEQ ID NO: 21**, the CDR-H2 comprises the sequence of **SEQ ID NO: 64**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 107**; 22) the CDR-H1 comprises the sequence of **SEQ ID NO: 22**, the CDR-H2 comprises the sequence of **SEQ ID NO: 65**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 108**; 23) the CDR-H1 comprises the sequence of **SEQ ID NO: 23**, the CDR-H2 comprises the sequence of **SEQ ID NO: 66**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 109**; 24) the CDR-H1 comprises the sequence of **SEQ ID NO: 24**, the CDR-H2 comprises the sequence of **SEQ ID NO: 67**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 110**; 25) the CDR-H1 comprises the sequence of **SEQ ID NO: 25**, the CDR-H2 comprises the sequence of **SEQ ID NO: 68**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 111**; 26) the CDR-H1 comprises the sequence of **SEQ ID NO: 26**, the CDR-H2 comprises the sequence of **SEQ ID NO: 69**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 112**; 27) the CDR-H1 comprises the sequence of **SEQ ID NO: 27**, the CDR-H2 comprises the sequence of **SEQ ID NO: 70**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 113**; 28) the CDR-H1 comprises the sequence of **SEQ ID NO: 28**, the CDR-H2 comprises the sequence of **SEQ ID NO: 71**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 114**; 29) the CDR-H1 comprises the sequence of **SEQ ID NO: 29**, the CDR-H2 comprises the sequence of **SEQ ID NO: 72**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 115**; 30) the CDR-H1 comprises the sequence of **SEQ ID NO: 30**, the CDR-H2 comprises the sequence of **SEQ ID NO: 73**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 116**; 31) the CDR-H1 comprises the sequence of **SEQ ID NO: 31**, the CDR-H2 comprises the sequence of **SEQ ID NO: 74**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 117**; 32) the CDR-H1 comprises the sequence of **SEQ ID NO: 32**, the CDR-H2 comprises the sequence of **SEQ ID NO: 75**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 118**; 33) the CDR-H1 comprises the sequence of **SEQ ID NO: 33**, the CDR-H2 comprises the sequence of **SEQ ID NO: 76**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 119**; 34) the CDR-H1 comprises the sequence of **SEQ ID NO: 34**, the CDR-H2 comprises the sequence of **SEQ ID NO: 77**, and the CDR-H3 comprises

the sequence of **SEQ ID NO: 120**; 35) the CDR-H1 comprises the sequence of **SEQ ID NO: 35**, the CDR-H2 comprises the sequence of **SEQ ID NO: 78**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 121**; 36) the CDR-H1 comprises the sequence of **SEQ ID NO: 36**, the CDR-H2 comprises the sequence of **SEQ ID NO: 79**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 122**; 37) the CDR-H1 comprises the sequence of **SEQ ID NO: 37**, the CDR-H2 comprises the sequence of **SEQ ID NO: 80**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 123**; 38) the CDR-H1 comprises the sequence of **SEQ ID NO: 38**, the CDR-H2 comprises the sequence of **SEQ ID NO: 81**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 124**; 39) the CDR-H1 comprises the sequence of **SEQ ID NO: 39**, the CDR-H2 comprises the sequence of **SEQ ID NO: 82**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 125**; 40) the CDR-H1 comprises the sequence of **SEQ ID NO: 40**, the CDR-H2 comprises the sequence of **SEQ ID NO: 83**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 126**; 41) the CDR-H1 comprises the sequence of **SEQ ID NO: 41**, the CDR-H2 comprises the sequence of **SEQ ID NO: 84**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 127**; 42) the CDR-H1 comprises the sequence of **SEQ ID NO: 42**, the CDR-H2 comprises the sequence of **SEQ ID NO: 85**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 128**; or 43) the CDR-H1 comprises the sequence of **SEQ ID NO: 43**, the CDR-H2 comprises the sequence of **SEQ ID NO: 86**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 129**.

[0074] In some embodiments, the CEA6 binding polypeptide comprise an immunoglobulin heavy chain variable domain comprising a CDR-H1, CDR-H2, and CDR-H3, where one or more of these CDRs are defined by a consensus sequence. The consensus sequences provided herein have been derived from the alignments of CDRs depicted in **FIG. 1**. However, it is envisioned that alternative alignments may be done (e.g. using global or local alignment, or with different algorithms, such as Hidden Markov Models, seeded guide trees, Needleman-Wunsch algorithm, or Smith-Waterman algorithm, or other known methods) and as such, alternative consensus sequences can be derived (including those done with a subset of the sequences provided herein).

[0075] In some embodiments, the CDR-H1 is defined by the formula $X_1X_2X_3X_4X_5X_6X_7X_8$, where X_1 is G; X_2 is F, R, S, or Y; X_3 is I or T; X_4 is F, G, L, S, or Y; X_5 is D, G, N, or S; X_6 is D, F, I, L, N, S, T, or Y; X_7 is D, N, or Y; X_8 is D, F, H, L, P, T, V, or Y. In some embodiments, the CDR-H1 comprises a sequence having at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to this consensus sequence. In some embodiments, the CDR-H1 comprises a sequence having 0, 1, 2, 3, 4, 5, or 6 substitutions from this consensus sequence.

[0076] In some embodiments, the CDR-H2 is defined by the formula $X_1X_2X_3X_4X_5X_6X_7X_8X_9X_{10}$, where X_1 is no amino acid, S, or T; X_2 is I; X_3 is N, S, or T; X_4 is R, S, T, or W; X_5 is D, F, I, L, S, T, or Y; X_6 is A, D, G, or S; X_7 is A, D, G, or S; X_8 is I or S; X_9 is T; X_{10} is no amino acid or Y. In some embodiments, the CDR-H2 comprises a sequence having at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to this consensus sequence. In some embodiments, the CDR-H2 comprises a sequence having 0, 1, 2, 3, 4, 5, or 6 substitutions from this consensus sequence.

[0077] In some embodiments, the CDR-H3 is defined by the formula $X_1X_2X_3X_4X_5X_6X_7X_8X_9X_{10}X_{11}X_{12}X_{13}X_{14}X_{15}X_{16}X_{17}X_{18}X_{19}X_{20}X_{21}X_{22}X_{23}X_{24}X_{25}X_{26}X_{27}X_{28}X_{29}X_{30}$ $X_{31}X_{32}X_{33}$, where X_1 is no amino acid or A; X_2 is no amino acid, A, or V; X_3 is no amino acid, A, G, M, Q, S, T, or V; X_4 is no amino acid, A, D, E, G, I, M, N, R, S, V, or Y; X_5 is no amino acid, A, E, K, M, R, S, T, V, or W; X_6 is no amino acid, A, E, M, P, S, or V; X_7 is no amino acid, A, F, I, M, P, W, or Y; X_8 is no amino acid, D, I, K, L, S, T, or V; X_9 is no amino acid, A, K, Q, T, or V; X_{10} is no amino acid, A, D, E, or S; X_{11} is no amino acid, A, I, L, R, V, or Y; X_{12} is no amino acid, A, E, G, L, P, S, or T; X_{13} is no amino acid, D, G, I, L, N, P, Q, S, T, or V; X_{14} is no amino acid, A, E, F, H, K, L, M, P, Q, R, T, V, or Y; X_{15} is no amino acid, A, D, H, I, L, M, P, Q, R, S, T, or V; X_{16} is no amino acid, A, D, E, H, L, S, T, V, W, or Y; X_{17} is no amino acid, E, F, G, H, L, M, N, Q, S, T, or Y; X_{18} is no amino acid, A, D, G, H, K, M, N, Q, R, S, V, or Y; X_{19} is no amino acid, F, H, Q, or Y; X_{20} is no amino acid, D, N, Q, S, or Y; X_{21} is no amino acid, A, G, or Y; X_{22} is no amino acid, W, or Y; X_{23} is no amino acid, A, or R; X_{24} is no amino acid, H, or S; X_{25} is no amino acid, D, or G; X_{26} is no amino acid, E, or K; X_{27} is no amino acid, I, or T; X_{28} is no amino acid, F, or R; X_{29} is no amino acid or Y; X_{30} is no amino acid or Y; X_{31} is no amino acid or Y; X_{32} is no amino acid, N, or S; X_{33} is no amino acid or Y.

[0078] In some embodiments, the CDR-H3 comprises a sequence having at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to this consensus sequence. In some embodiments, the CDR-H3 comprises a sequence having 0, 1, 2, 3, 4, 5, or 6 substitutions from this consensus sequence.

[0079] In some embodiments of the CEA6 binding polypeptides, the heavy chain variable domain comprises an amino acid sequence having at least 90%, 95%, 99%, or 100% sequence identity to any sequence selected from **SEQ ID NOs: 130-172**.

[0080] In some embodiments, the CEA6 binding polypeptide is humanized. In some embodiments, the CEA6 binding polypeptide is a single domain antibody (sdAb).

[0081] In some embodiments, the CEA6 binding polypeptide binds to CEA6 with a dissociation constant (KD) of less than 1 nM, 2 nM, 5 nM, 10 nM, 15 nM, 20 nM, 30 nM, 40 nM, 50 nM, 60 nM, 70 nM, 80 nM, 90 nM, 100 nM, 200 nM, 300 nM, 400 nM, 500 nM, 600 nM, 700

nM, 800 nM, 900 nM, or 1000 nM, or any KD within a range defined by any two of the aforementioned KD.

[0082] The binding polypeptides disclosed herein may be obtained from an antibody library. In some embodiments, the antibody library is an immune antibody library, a naïve antibody library, a synthetic antibody library, or a semi-synthetic antibody library. In some embodiments, the antibody library comprises antibodies derived from human, or antibodies that are not immunogenic in humans, or both. In some embodiments, the antibody library comprises antibodies that are humanized, e.g. from mouse, rat, guinea pig, rabbit, cat, dog, cow, horse, sheep, goat, horse, donkey. In some embodiments, the antibody library comprises single domain antibodies (sdAb), nanobodies, V_HH fragments, V_{NAR} fragments, single-chain variable fragments (scFv), camelid antibodies, or cartilaginous fish antibodies, or any combination thereof. One exemplary library that can be used is a fully humanized, synthetic, sdAb library, but any other antibody library that can be prepared or is available can be used for the methods disclosed herein. In some embodiments, the antibody library comprises sdAb. In some embodiments, the antibody library comprises at least 100, 500, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000, 20000, 30000, 40000, 50000, 60000, 70000, 80000, 90000, 100000, 500000, or 1000000 unique antibodies, or any number of antibodies within a range defined by any two of the aforementioned number of antibodies.

[0083] Antibody libraries may be generated computationally or using machine learning processes. An exemplary method of generating an antibody library computationally includes modifying a universal V_HH framework with synthetic diversity in one or more complementary determining regions (CDRs), such as CDR1, CDR2, or CDR3, or any combination thereof. The diversity of the CDRs are introduced by randomizing the library of sequences encoding for the antibodies with degenerate codons. For example, an NNK codon library can be employed, where the NNK codon comprises N (25% mix of A/T/C/G) and K (50% mix of T/G). In some embodiments, the NNK codon library is constructed with all possible amino acids, or with some amino acids (e.g. cysteine) and stop codon combinations excluded. Other degenerate codon mixes can be substituted for said NNK codon library with minimal experimentation. In other embodiments, the antibody library can be generated using a trimer codon mix, which improves balanced representation of sense codons while reducing the chance of stop codons, improving efficiency of antibody generation and testing. In some embodiments, artificial intelligence-based prediction can be used to randomize specific binding pockets of the antibodies using available binding models or structure data.

[0084] In some embodiments, panning the antibody library comprises screening for the candidate binding polypeptides by phage display, yeast display, bacterial display, ribosome

display, or mRNA display, or any combination thereof. In some embodiments, panning the antibody library comprises one or more rounds of selection, wherein the candidate binding polypeptides are selected for specificity towards a cancer-associated antigen (e.g. CEA6) or cells or tissues displaying the cancer-associated antigen. In some embodiments, the candidate binding polypeptides are selected under conditions including but not limited to tumor microenvironment-like conditions, immunosuppressive conditions, low or high pH, low or high oxygen concentrations, low or high temperatures, low or high viscosity, or any combination thereof, or for specificity towards modified or derivative forms of the one or more cancer-associated antigens. In some embodiments, the immunosuppressive conditions may comprise the presence of tumor-associated macrophages (TAMs), myeloid-derived suppressor cells (MDSCs), tumor-associated neutrophils (TANs), cancer-associated fibroblasts (CAFs), or other immunosuppressive cells, or the presence of adenosine, or both.

[0085] In some embodiments, the chimeric antigen receptor cells are from a cell line (e.g. Jurkat). In some embodiments, the chimeric antigen receptor cells are derived from a subject. In some embodiments, the subject has a cancer. In some embodiments, the subject has a cancer, and that cancer expresses any one or more of the cancer-associated antigens disclosed herein (e.g., CEA6). In some embodiments, the cancer is breast cancer, colorectal cancer, kidney cancer, liver cancer, lung cancer, brain cancer, pancreatic cancer, bladder cancer, testicular cancer, prostate cancer, gastric cancer, ovarian cancer, head and neck cancer, gallbladder cancer, a hematologic malignancy, or any combination thereof. In some embodiments, the hematologic malignancy may comprise leukemia, acute lymphoblastic leukemia, acute myelogenous leukemia, chronic myelogenous leukemia, chronic lymphocytic leukemia, hairy cell leukemia, lymphoma, Hodgkin's disease, Non-Hodgkin lymphoma, or multiple myeloma. In some embodiments, the subject is a mammal, such as a human, cat, dog, mouse, rat, hamster, rodent, cow, pig, horse, goat, sheep, donkey, or monkey. In some embodiments, the subject is a human.

Chimeric Antigen Receptors (CARs) and Cells

[0086] Also disclosed herein are chimeric antigen receptors (CARs) comprising any one or more of the CEA6 binding polypeptides disclosed herein.

[0087] In some embodiments, the CAR comprises at least two binding polypeptides and the CAR is a multivalent CAR. In some embodiments, the CAR comprises two binding polypeptides and the CAR is a bivalent CAR. In some embodiments, the CAR comprises three binding polypeptides and the CAR is a trivalent CAR.

[0088] In some embodiments, the CAR further comprises one or more signal peptides, linkers with various lengths and composition, hinges, transmembrane domains, costimulatory

domains, signaling domains, cytoplasmic domains, functionality signals, proliferation signals, anti-exhaustion signals, anti-inhibitory receptors, tumor/cancer homing proteins, or regulatory molecules, or any combination thereof. In some embodiments, the hinges comprise CD3ζ, CD4, CD8 or CD28 hinges, or computationally designed synthetic hinges with various lengths. In some embodiments, the transmembrane domains comprise CD3ζ, CD4, CD8 or CD28 transmembrane domains, or computationally designed synthetic transmembrane domains. In some embodiments, the costimulatory domains comprise CD8, CD28, ICOS, 4-1BB, OX40 (CD134), CD27, CD40, CD40L, TLR or other TNFR superfamily member or Ig superfamily member costimulatory domains, or other signaling via cytoplasmic domains of IL-2Rβ, IL-15R-α, MyD88, or CD40 or any other Toll-like receptor or IL-1 receptor signaling pathway members.

[0089] In some embodiments, the CARs disclosed herein are constructed by assembling CAR expression constructs from nucleic acids encoding for any one of the binding polypeptides disclosed herein and a mixture of compatible nucleic acids encoding for different CAR modules. In some embodiments, different combinations of CARs are produced for use in a CAR library for screening for CAR efficacy (*in vitro* or *in vivo*). In some embodiments, unique CARs are produced separately. In some embodiments, the CARs are specific for one target. In some embodiments, the CARs are specific for at least 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 targets. In some embodiments, the CARs are bi-specific or tri-specific.

[0090] To construct any one of CARs disclosed herein, the nucleic acids encoding for the binding polypeptides identified by panning of the antibody library are assembled into CAR expression constructs with other CAR modules. In some embodiments, the CAR expression constructs are assembled using multi-fragment assembly reactions known in the art. One exemplary method of assembling CAR expression constructs involves using Type IIS restriction enzymes to generate nucleic acid fragments with compatible overhang sequences and ligating the nucleic acid fragments with a ligase. As Type IIS restriction enzymes cleave outside of their recognition sites, multiple compatible nucleic acid fragments may be prepared simultaneously. In other embodiments, the CAR expression constructs can be assembled by overlap extension PCR. It is envisioned that any other method of assembling nucleic acid constructs from more than one nucleic acid fragment can be employed. In some embodiments, the different CAR modules comprise signal peptides, linkers, hinges, transmembrane domains, costimulatory domains, activation domains, signaling domains, cytoplasmic domains, functionality signals, proliferation signals, anti-exhaustion signals, anti-inhibitor receptors, cancer homing proteins, or regulatory molecules, or any combination thereof. Some exemplary hinges comprise CD8 hinge, CD28 hinge, IgG1 hinge, or IgG4 hinge. Some exemplary transmembrane domains comprise CD3ζ transmembrane domain, CD8α transmembrane domain, CD4 transmembrane domain, CD28

transmembrane domain, or ICOS transmembrane domain. Some exemplary costimulatory domains comprise CD8 costimulatory domain, CD28 costimulatory domain, 4-1BB costimulatory domain, OX40 (CD134) costimulatory domain, ICOS costimulatory domain, CD27 costimulatory domain, CD40 costimulatory domain, CD40L costimulatory domain, TLR costimulatory domains, MYD88-CD40 costimulatory domain, or KIR2DS2 costimulatory domain. In some embodiments, the different CAR modules are derived from CD8, CD28, 4-1BB, CD3ζ, or any combination thereof. The CAR may also be modified with various additions, including but not limited to cytokines, chemokines, cytokine receptors, chemokine receptors, antigen receptors or ligands, antibodies, or enzymes.

[0091] Also disclosed herein are chimeric antigen receptor (CAR) cells comprising any one of the CARs disclosed herein. In some embodiments, the CAR cell is a CAR-T cell. In some embodiments, the CAR cell is derived from a subject or from a cell line. In some embodiments, the subject has a cancer. In some embodiments, the subject has a cancer, and that cancer expresses any one or more of the cancer-associated antigens disclosed herein (e.g., CEA6). In some embodiments, the cancer is breast cancer, colorectal cancer, kidney cancer, liver cancer, lung cancer, brain cancer, pancreatic cancer, bladder cancer, testicular cancer, prostate cancer, gastric cancer, ovarian cancer, head and neck cancer, gallbladder cancer, a hematologic malignancy, or any combination thereof. In some embodiments, the hematologic malignancy may comprise leukemia, acute lymphoblastic leukemia, acute myelogenous leukemia, chronic myelogenous leukemia, chronic lymphocytic leukemia, hairy cell leukemia, lymphoma, Hodgkin's disease, Non-Hodgkin lymphoma, or multiple myeloma.

Nucleic Acids

[0092] Also disclosed herein are nucleic acids that encode for a polypeptide. In some embodiments, the polypeptide is a binding polypeptide. In some embodiments, the polypeptide is a single domain binding polypeptide. In some embodiments, the polypeptide is any one of the binding polypeptides disclosed herein. In some embodiments, the polypeptide comprises a sequence having at least 90%, 95%, 99%, or 100% sequence identity to any one or more of the CEA6 binding polypeptides disclosed herein. In some embodiments, the polypeptide is any one of the CARs disclosed herein.

[0093] Any one of the nucleic acids that encode for a binding polypeptide can be prepared by recombinant DNA technology, synthetic chemistry techniques, or a combination thereof. For example, sequences of nucleic acids encoding for the binding polypeptide may be cloned into an expression vector using standard molecular techniques known in the art. Sequences can be obtained from other vectors encoding the desired protein sequence, from PCR-generated

fragments using respective template nucleic acids, or by assembly of synthetic oligonucleotides encoding the desired sequences. In some embodiments, the expression vector may be a CAR expression vector, in which it is provided to an immune cell so that it expresses the CAR. In some embodiments, the expression vector may be an expression vector suited for large scale antibody or binding polypeptide production, from which the peptide products can be isolated for further use.

[0094] Expression of binding polypeptides or CARs may be confirmed by nucleic acid or protein assays known in the art. For example, the presence of transcribed mRNA of binding polypeptides or CARs can be detected and/or quantified by conventional hybridization assays (e.g. Northern blot analysis), amplification procedures (e.g. RT-PCR), SAGE (U.S. Pat. No. 5,695,937), and array-based technologies (see e.g. U.S. Pat. Nos. 5,405,783, 5,412,087 and 5,445,934), using probes complementary to any region of a polynucleotide that encodes for the binding polypeptides or CARs. Expression of the binding polypeptides or CARs can also be determined by examining the expressed peptide. A variety of techniques are available in the art for protein analysis. They include but are not limited to radioimmunoassays, ELISA (enzyme linked immunoradiometric assays), "sandwich" immunoassays, immunoradiometric assays, in situ immunoassays (using e.g., colloidal gold, enzyme or radioisotope labels), western blot analysis, immunoprecipitation assays, immunofluorescent assays, and SDS-PAGE.

Methods of Use or Treatment

[0095] Also disclosed herein are methods of treating a cancer in a subject in need thereof. In some embodiments, the methods comprise administering a chimeric antigen receptor cell to the subject. In some embodiments, the methods comprise administering any one of the chimeric antigen receptor cells disclosed herein. In some embodiments, the chimeric antigen receptor cell expresses and/or comprises any one or more of the CEA6 binding polypeptides disclosed herein. In some embodiments, the chimeric antigen receptor cell is a CAR-T cell. In some embodiments, the chimeric antigen receptor cell is derived from the subject and is autologous to the subject. In some embodiments, the chimeric antigen receptor cell is allogeneic to the subject. In some embodiments, the chimeric antigen receptor cell is from a cell line (e.g. Jurkat). In some embodiments, the subject is a mammal, such as a human, cat, dog, mouse, rat, hamster, rodent, cow, pig, horse, goat, sheep, donkey, or monkey. In some embodiments, the subject is a human. In some embodiments, the subject has a cancer, and that cancer expresses any one or more of the cancer-associated antigens disclosed herein (e.g., CEA6). In some embodiments, the cancer is breast cancer, colorectal cancer, kidney cancer, liver cancer, lung cancer, brain cancer, pancreatic cancer, bladder cancer, testicular cancer, prostate cancer, gastric

cancer, ovarian cancer, head and neck cancer, gallbladder cancer, a hematologic malignancy, or any combination thereof. In some embodiments, the hematologic malignancy may comprise leukemia, acute lymphoblastic leukemia, acute myelogenous leukemia, chronic myelogenous leukemia, chronic lymphocytic leukemia, hairy cell leukemia, lymphoma, Hodgkin's disease, Non-Hodgkin lymphoma, or multiple myeloma. In some embodiments, the chimeric antigen receptor cell is administered parenterally.

[0096] In some embodiments, the chimeric antigen receptor cell is administered once per day, twice per day, three times per day or more. In some embodiments, the chimeric antigen receptor cell is administered daily, every day, every alternate day, five days a week, once a week, every other week, two weeks per month, three weeks per month, once a month, twice a month, three times per month, or more. In some embodiments, the immune cell is administered for at least 1 month, 2 months, 3 months, 4 months, 5 months, 6 months, 7 months, 8 months, 9 months, 10 months, 11 months, 12 months, 18 months, 2 years, 3 years, or more.

[0097] In some embodiments, the amount of a given agent that correspond to such an amount varies depending upon factors such as the particular compound, the severity of the disease, the identity (e.g., weight) of the subject or host in need of treatment, but nevertheless is routinely determined in a manner known in the art according to the particular circumstances surrounding the case, including, e.g., the specific agent being administered, the route of administration, and the subject or host being treated. In some instances, the desired dose is conveniently presented in a single dose or as divided doses administered simultaneously (or over a short period of time) or at appropriate intervals, for example as two, three, four or more sub-doses per day.

[0098] The ranges for administration are merely suggestive, as the number of variables in regard to an individual treatment regime is large, and considerable excursions from these recommended values are not uncommon. Such dosages is altered depending on a number of variables, not limited to the activity of the compound used, the disease or condition to be treated, the mode of administration, the requirements of the individual subject, the severity of the disease or condition being treated, and the judgment of the practitioner.

[0099] In some embodiments, toxicity and therapeutic efficacy of such therapeutic regimens are determined by standard pharmaceutical procedures in cell cultures or experimental animals, including, but not limited to, the determination of the LD50 (the dose lethal to 50% of the population) and the ED50 (the dose therapeutically effective in 50% of the population). The dose ratio between the toxic and therapeutic effects is the therapeutic index and it is expressed as the ratio between LD50 and ED50. Compounds exhibiting high therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used in formulating a range of dosage for use in human. The dosage of such compounds lies preferably within a range of

circulating concentrations that include the ED50 with minimal toxicity. The dosage varies within this range depending upon the dosage form employed and the route of administration utilized.

EXAMPLES

[0100] Some aspects of the embodiments discussed above are disclosed in further detail in the following examples, which are not in any way intended to limit the scope of the present disclosure. Those in the art will appreciate that many other embodiments also fall within the scope of the invention, as it is described herein above and in the claims.

Example 1. Methodology

[0101] Plasmid preparation and quality check (QC)

[0102] The Maxi Plasmid Purification kit (Zymo, D4203) was used for CAR plasmid preparation. Plasmid concentration and quality was analyzed by Nanodrop (260/280 ratio and 260/230 ratio) and the ToxinSensor Chromogenic LAL Endotoxin Assay (Genescript, L00350). Good-quality plasmid DNA will have an A260/A280 ratio of 1.8-2.0, an A260/A230 ratio greater than 2.0, and less than 0.1 EU/μg of endotoxin.

[0103] Plasmid assembly and QC

[0104] All CARs were synthesized by Genewiz. Synthetic single domain antibody genes encoding antibodies which bind to CEA6 were cloned into the pLenti vector and the construct was confirmed using Sanger sequencing. The CAR constructs contain a signal peptide (e.g. CD8 signal peptide), hinge (e.g. CD8α stalk), transmembrane domain (e.g. CD8 transmembrane domain or CD28 transmembrane domain), and signaling domains (e.g. 4-1BB costimulatory domain or CD3ζ signaling domain) and the anti-CEA6 sdAb.

[0105] To monitor CAR expression levels, the expressed cassette of the CAR plasmids can also be engineered to express fluorescent eGFP+T2A self-cleaving peptide sequences. Alternatively, truncated CD19 or truncated EGFR cassettes can be used to monitor by antibody detection. Human sequences (e.g. CD8, 4-1BB, CD3ζ) are accessible on GenBank.

[0106] Virus production and titration

[0107] To produce the lentivirus with the anti-CEA6 CAR construct, human embryonic kidney 293T (HEK293T) cells were co-transfected with CAR transgene-encoding pLenti transfer plasmid and one or more necessary packaging plasmids (i.e. encoding for Gag, Pol, Rev, VSVG, and optionally Tat). The supernatants of the HEK293T cultures were collected at either 48 or 72 hours after transfection, centrifuged, and filtered with a 0.45 μm filter.

[0108] Virus titration was done in Jurkat cells transduced with diluted lentivirus collections. After 48 hours, transduced Jurkat cells were stained with biotinylated recombinant

protein L and phycoerythrin (PE)-conjugated streptavidin, and anti-CEA6 CAR abundance was measured by flow cytometry.

[0109] T cell transduction and expansion

[0110] Human PBMCs were isolated from peripheral blood of healthy human donors by density gradient centrifugation with the Lymphoprep reagent (StemCell Technologies). PBMCs were resuspended at $1x10^6$ cells/mL in X-VIVO 15 serum-free hematopoietic medium (Lonza, 04-418QCN) with 10 ng/mL IL-2 (Novoprotein, GMP-CD66) and 10 ng/mL IL-7 (Novoprotein, GMP-CD47). PBMCs were stimulated with 50 ng/mL anti-CD3 antibody (Novoprotein, GMP-A018) for 24 hours. Then, PBMCs were transduced with anti-CEA6 CAR lentivirus (at MOI of 1-10) and expanded in appropriate flasks for 9-10 days in RPMI 1640 basal medium supplemented with 10% fetal bovine serum. T cells were rested for 24 hours at 37°C before any assay. CAR surface levels, CD3, and CD4/CD8 ratios (using antibodies from Biolegend, 317344, 301012, and 317412) were measured 12 or 14 days after initial stimulation of PBMCs with the anti-CD3 antibody.

[0111] T cell cryopreservation and thawing

[0112] After 14 days of CAR-T expansion, anti-CEA6 CAR-T cells were centrifuged and the supernatant was discarded. The cell pellet was resuspended in chilled CryoStor CS10 (StemCell Technologies, 07930) at a viable cell density of 5×10^7 cells/mL. Aliquots of cell suspension were dispensed into cryovials. The cryovials were cooled with a 1°C/minute decrease. The frozen cells were transferred to liquid nitrogen.

[0113] To thaw, the cells are quickly thawed in a 37°C water bath with gentle agitation. The thawed cells are transferred to a 50 mL conical tube and washed by adding 20 mL of fresh growth medium dropwise. The cells are centrifuged and resuspended in X-VIVO 15 medium at a cell density of 1×10^6 cells/mL.

[0114] Cytotoxicity assay

[0115] The cytotoxicity of anti-CEA6 CAR T-cells was determined by standard luciferase-based assays. Briefly, target cells expressing firefly luciferase were co-cultured with CAR-T cells in triplicate at the indicated effector:target ratios using white-walled 96-well plates with 2×10^4 target cells in a total volume of 100 µL per well in X-VIVO 15 medium. Target cells alone as a control were plated at the same cell density. After 48 hours of co-culture, 100 µL of luciferase substrate (ONE-Glo, Promega) was directly added to each well. Emitted light was detected with a luminescence plate reader.

[0116] In a similar luciferase based cytotoxicity assay, target cancer cells engineered to overexpress luciferase were prepared in a cell culture medium (such as RPMI) and seeded on to 96 well plates with 5×10^4 to 5×10^6 cells in a volume of 100 µL/well, and incubated at 37°C for

24 hours. Then, CAR T-cells were seeded in corresponding wells with CAR T-cells at various effector:target (E:T) ratios (e.g., 5:1, 1:1, and 1:5) in 100 µL/well. The co-cultures were then incubated at 37°C for 24 or 48 hours. Subsequently, the cells were processed with the Luciferase Assay System (Promega, E1501) according to manufacturer's instructions using a plate reader. Luminescence was read in endpoint mode using a BioTek Synergy H4 hybrid microplate reader for 10 seconds. Percentage of specific lysis was calculated using the luciferase activity of digitonin-treated cells as maximum cell death and untreated cells as spontaneous cell death, using the formula % specific lysis = 100% x [(experimental luminescence – spontaneous cell death luminescence)/(maximum cell death luminescence – spontaneous cell death luminescence)].

[0117] Additional information regarding exemplary luciferase cytotoxicity assays may be found in Matta H et al. "Development and characterization of a novel luciferase based cytotoxicity assay" *Scientific Reports* (2018) 8,199, which is hereby expressly incorporated by reference in its entirety.

[0118] IFN-γ evaluation

[0119] 1 x 10^6 per well of anti-CEA6 CAR T-cells were stimulated with target cells at indicated effector:target cell ratios in 6-well tissue culture plates for 24 hours. Interferon gamma (IFN-γ) and IL-2 secretion was quantitated through enzyme-linked immunosorbent assay (ELISA) using the human IL-2 ELISA Kit II (BD, 550611) and human IFN-γ ELISA Kit II (BD, 550612)

[0120] CD25/CD69 assay

[0121] 1 x 10^6 per well of anti-CEA6 CAR T-cells were stimulated with target cells at indicated effector:target cell ratios in 6-well tissue culture plates for 24 hours. Cells were stained with LIVE/DEAD Fixable Near-IR Dead Cell Stain Kit (ThermoFisher) to label dead cells and with anti-CD3, and CAR-T cells were identified as CD3+ GFP+ cells by flow cytometry. CD69 and CD25 on CAR-T cells were stained with Alexa Fluor 700 anti-human CD69 antibody (Biolegend) and PE anti-CD25 antibody (Biolegend).

[0122] Sequencing of CAR T cell libraries

[0123] Cells were collected from anti-CEA6 CAR-T library screens. Genomic DNA (gDNA) was extracted using the Blood/Cell Genome DNA Mini kit (Tiangen, DP304-03). 75 ng of gDNA was used for PCT to amplify the CAR region of the library. The amplified region was TA cloned with the TA/Blunt-Zero Cloning kit (Vazyme, C601) and sequenced.

Example 2. Exemplary anti-CEA6 CAR T-cell lines are effective against cancers

[0124] Several anti-CEA6 CAR T-cell lines were prepared from the anti-CEA6 binders disclosed herein. Constructs were expressed in T-cells isolated from two donors, denoted Z0016 and Y1287. Cytotoxicity assays were performed as described in Example 1, at effector:target

ratios of 1:5, 1:1, and 5:1. These anti-CEA6 CAR T-cell lines showed cytotoxic efficacy against MKN-45 gastric cancer cells (**FIG. 2A-B**), BxPC-3 pancreatic cancer cells (**FIG. 2C-D**), and Capan-1 pancreatic cancer cells (**FIG. 2E-F**). Untransfected T-cells (UNT) were used as negative control.

[0125] An exemplary anti-CEA6 CAR T-cell line B4T2-001 was tested in *in vivo* NSG mouse models. In a cell line-derived xenograft model, administration of B4T2-001 resulted in dramatic clearance of the xenograft for gastric, non-small cell lung cancer, and pancreatic tumors whereas tumor growth in control mice progressed (**FIG. 3A**). In an orthotopic pancreatic cancer model, robust clearance of the pancreatic tumor (as detected by luciferase luminescence) resulted from 23 days of B4T2-001 treatment, whereas control mice exhibited no reduction in tumor size (**FIG. 3B**). In a patient-derived xenograph (PDX) of colorectal cancer, mice treated with B4T2-001 exhibited clearance of the tumor as early as 42 days.

[0126] A tumor re-challenge study using the anti-CEA6 CAR T-cell line B4T2-001 was performed and the results are shown in **FIG. 4**. An initial CAR T-cell infusion was administered to a pancreatic tumor model at day 0. While control mice exhibited unrestrained tumor growth, mice administered B4T2-001 exhibited undetectable tumor volume as early as day 19. At day 87, where the presence of B4T2-001 T-cells was detected to be approximately 100 cells per 100 μL of blood, a second tumor was engrafted into the same mice. A rapid re-expansion of the anti-CEA6 CAR T-cells was observed, and the re-engrafted tumor, which reached a peak volume at approximately day 107, was completely cleared at around day 120. This shows that B4T2-001 remains persistent for a significant amount of time and is capable of driving a secondary complete response.

Example 3. Anti-CEA6 CAR-T cells for the treatment of cancer

[0127] A patient presents with a cancer, for example, a cancer that expresses or overexpresses CEA6 (e.g., such that CEA6 can be used as a biomarker to detect and target the cancer). In some cases, the cancer is breast cancer, colorectal cancer, kidney cancer, liver cancer, lung cancer, brain cancer, pancreatic cancer, bladder cancer, testicular cancer, prostate cancer, gastric cancer, ovarian cancer, head and neck cancer, gallbladder cancer, a hematologic malignancy, or any combination thereof. In some cases, the hematologic malignancy may comprise leukemia, acute lymphoblastic leukemia, acute myelogenous leukemia, chronic myelogenous leukemia, chronic lymphocytic leukemia, hairy cell leukemia, lymphoma, Hodgkin's disease, Non-Hodgkin lymphoma, or multiple myeloma.

[0128] One or more of the CAR T-cells disclosed herein, which may include any one or more of the anti-CEA6 CARs and/or anti-CEA6 binders, are administered to the patient enterally,

orally, intranasally, parenterally, intravenously, intraperitoneally, intramuscularly, intra-arterially, intraventricularly, intradermally, intralesionally, intracranially, intrathecally, or subcutaneously.

[0129] The CAR T-cells are administered as doses in the amount of 10^4, 10^5, 10^6, 10^7, 10^8, or 10^9 cells per dose, or any number of cells within a range defined by any two of the aforementioned cells per dose, or any number of cells that is effective and/or safe as determined by a trained medical practitioner. The doses are administered every 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 24, 36, or 48 days, or weeks, or any time within a range defined by any two of the aforementioned times, or any timing of dosing that is effective and/or safe as determined by a trained medical practitioner.

[0130] An improvement of the cancer or symptoms thereof is observed in the patient following administration of the one or more CAR T-cells. Administration of the CAR T-cells may be performed in conjunction with another therapy for cancer, including but not limited to immunotherapy, chemotherapy, radiation therapy, surgery, photodynamic therapy, or targeted therapy.

[0131] In at least some of the previously described embodiments, one or more elements used in an embodiment can interchangeably be used in another embodiment unless such a replacement is not technically feasible. It will be appreciated by those skilled in the art that various other omissions, additions and modifications may be made to the methods and structures described above without departing from the scope of the claimed subject matter. All such modifications and changes are intended to fall within the scope of the subject matter, as defined by the appended claims.

[0132] With respect to the use of substantially any plural and/or singular terms herein, those having skill in the art can translate from the plural to the singular and/or from the singular to the plural as is appropriate to the context and/or application. The various singular/plural permutations may be expressly set forth herein for sake of clarity.

[0133] It will be understood by those within the art that, in general, terms used herein, and especially in the appended claims (e.g., bodies of the appended claims) are generally intended as "open" terms (e.g., the term "including" should be interpreted as "including but not limited to," the term "having" should be interpreted as "having at least," the term "includes" should be interpreted as "includes but is not limited to," etc.). It will be further understood by those within the art that if a specific number of an introduced claim recitation is intended, such an intent will be explicitly recited in the claim, and in the absence of such recitation no such intent is present. For example, as an aid to understanding, the following appended claims may contain usage of the introductory phrases "at least one" and "one or more" to introduce claim recitations. However, the

use of such phrases should not be construed to imply that the introduction of a claim recitation by the indefinite articles "a" or "an" limits any particular claim containing such introduced claim recitation to embodiments containing only one such recitation, even when the same claim includes the introductory phrases "one or more" or "at least one" and indefinite articles such as "a" or "an" (e.g., "a" and/or "an" should be interpreted to mean "at least one" or "one or more"); the same holds true for the use of definite articles used to introduce claim recitations. In addition, even if a specific number of an introduced claim recitation is explicitly recited, those skilled in the art will recognize that such recitation should be interpreted to mean at least the recited number (e.g., the bare recitation of "two recitations," without other modifiers, means at least two recitations, or two or more recitations). Furthermore, in those instances where a convention analogous to "at least one of A, B, and C, etc." is used, in general such a construction is intended in the sense one having skill in the art would understand the convention (e.g., "a system having at least one of A, B, and C" would include but not be limited to systems that have A alone, B alone, C alone, A and B together, A and C together, B and C together, and/or A, B, and C together, etc.). In those instances where a convention analogous to "at least one of A, B, or C, etc." is used, in general such a construction is intended in the sense one having skill in the art would understand the convention (e.g., " a system having at least one of A, B, or C" would include but not be limited to systems that have A alone, B alone, C alone, A and B together, A and C together, B and C together, and/or A, B, and C together, etc.). It will be further understood by those within the art that virtually any disjunctive word and/or phrase presenting two or more alternative terms, whether in the description, claims, or drawings, should be understood to contemplate the possibilities of including one of the terms, either of the terms, or both terms. For example, the phrase "A or B" will be understood to include the possibilities of "A" or "B" or "A and B."

[0134] In addition, where features or aspects of the disclosure are described in terms of Markush groups, those skilled in the art will recognize that the disclosure is also thereby described in terms of any individual member or subgroup of members of the Markush group.

[0135] As will be understood by one skilled in the art, for any and all purposes, such as in terms of providing a written description, all ranges disclosed herein also encompass any and all possible sub-ranges and combinations of sub-ranges thereof. Any listed range can be easily recognized as sufficiently describing and enabling the same range being broken down into at least equal halves, thirds, quarters, fifths, tenths, etc. As a non-limiting example, each range discussed herein can be readily broken down into a lower third, middle third and upper third, etc. As will also be understood by one skilled in the art all language such as "up to," "at least," "greater than," "less than," and the like include the number recited and refer to ranges which can be subsequently broken down into sub-ranges as discussed above. Finally, as will be understood by one skilled in

the art, a range includes each individual member. Thus, for example, a group having 1-3 articles refers to groups having 1, 2, or 3 articles. Similarly, a group having 1-5 articles refers to groups having 1, 2, 3, 4, or 5 articles, and so forth.

[0136] While various aspects and embodiments have been disclosed herein, other aspects and embodiments will be apparent to those skilled in the art. The various aspects and embodiments disclosed herein are for purposes of illustration and are not intended to be limiting, with the true scope and spirit being indicated by the following claims.

[0137] All references cited herein, including but not limited to published and unpublished applications, patents, and literature references, are incorporated herein by reference in their entirety and are hereby made a part of this specification. To the extent publications and patents or patent applications incorporated by reference contradict the disclosure contained in the specification, the specification is intended to supersede and/or take precedence over any such contradictory material.

Sequence List

CEA6 sequences

Name	CDR-H1	SEQ ID NO	CDR-H2	SEQ ID NO	CDR-H3	SEQ ID NO
CEA6-H01	GSIYSYYH	1	ISRTGGST	44	AAARPMITHSLMYYDY	87
CEA6-F02	GFIYSFYF	2	ISRTGGST	45	AGIRMITIEHWNKYSY	88
CEA6-F06	GSISGSYY	3	TINSFGDITY	46	ANTGHSDFSY	89
CEA6-H06	GSISSFYV	4	ISRTGGST	47	AAKTALTWYNYDY	90
CEA6-B07	GSIYSFDF	5	ISRTGGST	48	AAKTALSILDLTAFDY	91
CEA6-A08	GYILSFNF	6	ISRTGGST	49	ASAPGPSLTVFDY	92
CEA6-F09	GSIFSFNT	7	TISWSSDSTY	50	ANTGHSASSHNY	93
CEA6-B11	GSISSSNF	8	ISRTGGST	51	VAMENHITLDFDY	94
CEA6-D11	GSILSFYV	9	ISRTGGST	52	VYQAWAHGKIFNY	95
Cap01+02-CEA6-Bio-R2-3	GSIFDLYY	10	TINSLDAITY	53	AANTGYSSLSY	96
Cap01+02-CEA6-Bio-R2-11	GRISDIYY	11	SINSYADITY	54	AANTGFSSLSY	97
Cap03+04-CEA6-R2-1	GFIFDNYP	12	ITSDGDST	55	AVLPFPWNMQYGY	98
Cap03+04-CEA6-R2-2/5	GSISNIYD	13	ISRTGGST	56	AASEVPPLTLAEQYYDY	99
Cap03+04-CEA6-R2-4	GSIYNFYL	14	ISSFGGIT	57	AAWPYVQDRATLLAHNY	100
Cap03+04-CEA6-R2-7	GRTFSYNP	15	ISRTGGST	58	AAIEDVLWHNY	101
Cap03+04-CEA6-R2-9	GRIYNTDH	16	ISRTGGST	59	AAQDKVRPTLKFDY	102
Cap03+04-CEA6-R2-10	GRIYSFNF	17	ISRTGGST	60	AAGVMVYVVTDHQRYDY	103
Cap03+04-CEA6-R2-11/8	GFTFSTNF	18	INSFGSST	61	AVVVSFSTEYGSYVYYGYDY	104
Cap03+04-CEA6-R2-13/6	GSIYSFYF	19	ISRTGGST	62	AVTRTAITLEQVQQYSY	105
Cap03+04-CEA6-R2-15	GRISGFYF	20	INTIAAST	63	AAKEWDKDATTYSY	106
Cap03+04-CEA6-R2-16	GSIFDFYF	21	ISRTGGST	64	AATYSPPIQMMSHHFDY	107
Cap03+04-CEA6-R2-19	GRTFGLNF	22	ISRTGGST	65	AAMEKSKPSLDFDY	108
Cap03+04-CEA6-R3-12	GSISNIYF	23	INSISDST	66	AVRRWGGYDAYRSDETRYYYSY	109
Cap03+04-CEA6-R3-15	GSIYGFYF	24	ISRTGGST	67	AASRALVSVQQHEGFNY	110
Cap03+04-CEA6-R3-39/19	GSIFNFYF	25	ISRTGGST	68	AASGATVDVADASSYNY	111
Cap03+04-CEA6-R3-1	GFIYSFYF	26	ISRTGGST	69	AAGIRMITIEHWNKYSY	112
Cap03+04-CEA6-R3-11	GYILSFNF	27	ISRTGGST	70	AASAPGPSLTVFDY	113

Name	CDR-H1	SEQ ID NO	CDR-H2	SEQ ID NO	CDR-H3	SEQ ID NO
Cap01+02-CEA6-Bio-R2-1	GSISGSYY	28	TINSFGDITY	71	AANTGHSDFSY	114
Cap01+02-CEA6-R3-32	GSIFSFNT	29	TISWSSDSTY	72	AANTGHSASSHNY	115
Cap03+04-CEA6-R2-17/3	GSISSFYV	30	ISRTGGST	73	AAAKTALTWYNYDY	116
CEA6-cell-R3-65	GSISGSYY	31	TINSFGDITY	74	AANTGHSDFSY	117
CEA6-cell-R3-66	GSIFSFNT	32	TISWSSDSTY	75	AANTGHSASSHNY	118
H002-10-41BB	GSIYSFYF	33	ISRTGGST	76	AVTRTAITLEQVQQYSY	119
H002-2-41BB	GRISDIYY	34	SINSYADITY	77	ANTGFSSLSY	120
H002-3-41BB	GFIFDNYP	35	ITSDGDST	78	AVLPFPWNMQYGY	121
H002-3-BC141BB	GSISGSYY	36	TINSFGDITY	79	ANTGHSDFSY	122
H002-6-41BB	GRTFSYNP	37	ISRTGGST	80	AIEDVLWHNY	123
H002-7-BC141BB	GSIFSFNT	38	TISWSSDSTY	81	ANTGHSASSHNY	124
H002-8-41BB	GRIYSFNF	39	ISRTGGST	82	AGVMVYVVTDHQRYDY	125
H002-8-CP3-41BB	GFIGNDYD	40	ISRTGGST	83	AASTTFSSHNY	126
H002-M-1	GFIYSFYF	41	ISRTGGST	84	AGIRMITIEHWNKYSY	127
H002-M-2	GFIYSFYF	42	ISRTGGST	85	ANTGHSDFSY	128
H002-M-3	GSISGSYY	43	TINSFGDITY	86	AGIRMITIEHWNKYSY	129

Name	Heavy Chain Variable Domain Sequence	SEQ ID NO:
CEA6-H01	EVQLVESGGGLVQPGGSLRLSCAASGSIYSYYHMGWFRQAPGKGRELVAAISRTGGS TYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCAAAARPMITHSLMYYDYW GQGTQVTVSS	130
CEA6-F02	EVQLVESGGGLVQPGGSLRLSCAASGFIYSFYFMGWFRQAPGKGRELVAAISRTGGS TYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCAAGIRMITIEHWNKYSYW GQGTQVTVSS	131
CEA6-F06	EVQLVESGGGLVQPGGSLRLSCAASGSISGSYYMGWFRQAPGKGRELVAATINSFGD ITYYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCAANTGHSDFSYWGQGT QVTVSS	132
CEA6-H06	EVQLVESGGGLVQPGGSLRLSCAASGSISSFYVMGWFRQAPGKGRELVAAISRTGGS TYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCAAAKTALTWYNYDYWGQG TQVTVSS	133
CEA6-B07	EVQLVESGGGLVQPGGSLRLSCAASGSIYSFDFMGWFRQAPGKGRELVAAISRTGGS TYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCAAAKTALSILDLTAFDYW GQGTQVTVSS	134
CEA6-A08	EVQLVESGGGLVQPGGSLRLSCAASGYILSFNFMGWFRQAPGKGRELVAAISRTGGS TYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCAASAPGPSLTVFDYWGQG TQVTVSS	135
CEA6-F09	EVQLVESGGGLVQPGGSLRLSCAASGSIFSFNTMGWFRQAPGKGRELVAATISWSSD STYYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCAANTGHSASSHNYWGQ GTQVTVSS	136

Name	Heavy Chain Variable Domain Sequence	SEQ ID NO:
CEA6-B11	EVQLVESGGGLVQPGGSLRLSCAASGSISSSNFMGWFRQAPGKGRELVAAISRTGGS TYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCAVAMENHITLDFDYWGQG TQVTVSS	137
CEA6-D11	EVQLVESGGGLVQPGGSLRLSCAASGSILSFYVMGWFRQAPGKGRELVAAISRTGGS TYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCAVYQAWAHGKIFNYWGQG TQVTVSS	138
Cap01+02-CEA6-Bio-R2-3	EVQLVESGGGLVQPGGSLRLSCAASGSIFDLYYMGWFRQAPGKGRELVAATINSLDA ITYYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCAANTGYSSLSYWGRGT QVTVSS	139
Cap01+02-CEA6-Bio-R2-11	EVQLVESGGGLVQPGGSLRLSCAASGRISDIYYMGWFRQAPGKGRELVAASINSYAD ITYYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCAANTGFSSLSYWGQGT QVTVSS	140
Cap03+04-CEA6-R2-1	EVQLVESGGGLVQPGGSLRLSCAASGFIFDNYPMGWFRQAPGKGRELVATITSDGDS TYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCAVLPFPWNMQYGYWGQGT QVTVSS	141
Cap03+04-CEA6-R2-2/5	EVQLVESGGGLVQPGGSLRLSCAASGSISNIYDMGWFRQAPGKGRELVAAISRTGGS TYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCAASEVPPLTLAEQYYDYW GQGTQVTVSS	142
Cap03+04-CEA6-R2-4	EVQLVESGGGLVQPGGSLRLSCAASGSIYNFYLMGWFRQAPGKGRELVATISSFGGI TYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCAAWPYVQDRATLLAHNYW GQGTQVTVSS	143
Cap03+04-CEA6-R2-7	EVQLVESGGGLVQPGGSLRLSCAASGRTFSYNPMGWFRQAPGKGRELVAAISRTGGS TYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCAAIEDVLWHNYWGQGTQV TVSS	144
Cap03+04-CEA6-R2-9	EVQLVESGGGLVQPGGSLRLSCAASGRIYNTDHMGWFRQAPGKGRELVAAISRTGGS TYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCAAQDKVRPTLKFDYWGQG TQVTVSS	145
Cap03+04-CEA6-R2-10	EVQLVESGGGLVQPGGSLRLSCAASGRIYSFNFMGWFRQAPGKGRELVAAISRTGGS TYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCAAGVMVYVVTDHQRYDYW GQGTQVTVSS	146
Cap03+04-CEA6-R2-11/8	EVQLVESGGGLVQPGGSLRLSCAASGFTFSTNFMGWFRQAPGKGRELVATINSFGSS TYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCAVVVSFSTEYGSYVYYGY DYWGQGTQVTQVTVSS	147
Cap03+04-CEA6-R2-13/6	EVQLVESGGGLVQPGGSLRLSCAASGSIYSFYFMGWFRQAPGKGRELVAAISRTGGS TYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCAVTRTAITLEQVQQYSYW GQGTQVTVSS	148
Cap03+04-CEA6-R2-15	EVQLVESGGGLVQPGGSLRLSCAASGRISGFYFMGWFRQAPGKGRELVAAINTIAAS TYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCAAKEWDKDATTYSYWGQG TQVTVSS	149
Cap03+04-CEA6-R2-16	EVQLVESGGGLVQPGGSLRLSCAASGSIFDFYFMGWFRQAPGKGRELVAAISRTGGS TYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCAATYSPPIQMMSHHFDYW GQGTQVTVSS	150
Cap03+04-CEA6-R2-19	EVQLVESGGGLVQPGGSLRLSCAASGRTFGLNFMGWFRQAPGKGRELVAAISRTGGS TYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCAAMEKSKPSLDFDYWGQG TQVTVSS	151
Cap03+04-CEA6-R3-12	EVQLVESGGGLVQPGGSLRLSCAASGSISNIYFMGWFRQAPGKGRELVAAINSISDS TYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCAVRRWGGYDAYRSDETRY YYSYWGQGTQVTVSS	152
Cap03+04-CEA6-R3-15	EVQLVESGGGLVQPGGSLRLSCAASGSIYGFYFMGWFRQAPGKGRELVAAISRTGGS TYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCAASRALVSVQQHEGFNYW GQGTQVTVSS	153
Cap03+04-CEA6-R3-39/19	EVQLVESGGGLVQPGGSLRLSCAASGSIFNFYFMGWFRQAPGKGRELVAAISRTGGS TYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCAASGATVDVADASSYNYW GQGTQVTVSC	154
Cap03+04-CEA6-R3-1	EVQLVESGGGLVQPGGSLRLSCAASGFIYSFYFMGWFRQAPGKGRELVAAISRTGGS TYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCAAGIRMITIEHWNKYSYW GQGTQVTVSS	155
Cap03+04-CEA6-R3-11	EVQLVESGGGLVQPGGSLRLSCAASGYILSFNFMGWFRQAPGKGRELVAAISRTGGS TYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCAASAPGPSLTVFDYWGQG TQVTVSS	156

Name	Heavy Chain Variable Domain Sequence	SEQ ID NO:
Cap01+02-CEA6-Bio-R2-1	EVQLVESGGGLVQPGGSLRLSCAASGSISGSYYMGWFRQAPGKGRELVAATINSFGD ITYYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCAANTGHSDFSYWGQGT QVTVSS	157
Cap01+02-CEA6-R3-32	EVQLVESGGGLVQPGGSLRLSCAASGSIFSFNTMGWFRQAPGKGRELVAATISWSSD STYYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCAANTGHSASSHNYWGQ GTQVTVSS	158
Cap03+04-CEA6-R2-17/3	EVQLVESGGGLVQPGGSLRLSCAASGSISSFYVMGWFRQAPGKGRELVAAISRTGGS TYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCAAAKTALTWYNYDYWGQG TQVTVSS	159
CEA6-cell-R3-65	EVQLVESGGGLVQPGGSLRLSCAASGSISGSYYMGWFRQAPGKGRELVAATINSFGD ITYYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCAANTGHSDFSYWGQGT QVTVSS	160
CEA6-cell-R3-66	EVQLVESGGGLVQPGGSLRLSCAASGSIFSFNTMGWFRQAPGKGRELVAATISWSSD STYYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCAANTGHSASSHNYWGQ GTQVTVSS	161
H002-10-41BB	EVQLVESGGGLVQPGGSLRLSCAASGSIYSFYFMGWFRQAPGKGRELVAAISRTGGS TYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCAVTRTAITLEQVQQYSYW GQGTQVTVSS	162
H002-2-41BB	EVQLVESGGGLVQPGGSLRLSCAASGRISDIYYMGWFRQAPGKGRELVAASINSYAD ITYYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCANTGFSSLSYWGQGTQ VTVSS	163
H002-3-41BB	EVQLVESGGGLVQPGGSLRLSCAASGFIFDNYPMGWFRQAPGKGRELVAAITSDGDS TYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCAVLPFPWNMQYGYWGQGT QVTVSS	164
H002-3-BC141BB	EVQLVESGGGLVQPGGSLRLSCAASGSISGSYYMGWFRQAPGKGRELVAATINSFGD ITYYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCANTGHSDFSYWGQGTQ VTVSS	165
H002-6-41BB	EVQLVESGGGLVQPGGSLRLSCAASGRTFSYNPMGWFRQAPGKGRELVAAISRTGGS TYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCAIEDVLWHNYWGQGTQVT VSS	166
H002-7-BC141BB	EVQLVESGGGLVQPGGSLRLSCAASGSIFSFNTMGWFRQAPGKGRELVAATISWSSD STYYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCANTGHSASSHNYWGQG TQVTVSS	167
H002-8-41BB	EVQLVESGGGLVQPGGSLRLSCAASGRIYSFNFMGWFRQAPGKGRELVAAISRTGGS TYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCAGVMVYVVTDHQRYDYWG QGTQVTVSS	168
H002-8-CP3-41BB	EVQLVESGGGLVQPGGSLRLSCAASGFIGNDYDMGWFRQAPGKGRELVAAISRTGGS TYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCAASTTFSSHNYWGQGTQV TVSS	169
H002-M-1	EVQLVESGGGLVQPGGSLRLSCAASGFIYSFYFMGWFRQAPGKGRELVAAISRTGGS TYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCAGIRMITIEHWNKYSYWG QGTQVTVSS	170
H002-M-2	EVQLVESGGGLVQPGGSLRLSCAASGFIYSFYFMGWFRQAPGKGRELVAAISRTGGS TYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCANTGHSDFSYWGQGTQVT VSS	171
H002-M-3	EVQLVESGGGLVQPGGSLRLSCAASGSISGSYYMGWFRQAPGKGRELVAATINSFGD ITYYYPDSVEGRFTISRDNAKRMVYLQMNSLRAEDTAVYYCAGIRMITIEHWNKYSY WGQGTQVTVSS	172

WHAT IS CLAIMED IS:

1. A carcinoembryonic antigen 6 (CEA6) binding polypeptide comprising an immunoglobulin heavy chain variable domain comprising a CDR-H1, CDR-H2, and CDR-H3, wherein:

the CDR-H1 comprises a sequence having at least 90%, 95%, 99%, or 100% sequence identity to a sequence selected from **SEQ ID NOs: 1-43**;

the CDR-H2 comprises a sequence having at least 90%, 95%, 99%, or 100% sequence identity to a sequence selected from **SEQ ID NOs: 44-86**; and

the CDR-H3 comprises a sequence having at least 90%, 95%, 99%, or 100% sequence identity to a sequence selected from **SEQ ID NOs: 87-129**.

2. The CEA6 binding polypeptide of claim 1, wherein:

1) the CDR-H1 comprises the sequence of **SEQ ID NO: 1**, the CDR-H2 comprises the sequence of **SEQ ID NO: 44**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 87**;

2) the CDR-H1 comprises the sequence of **SEQ ID NO: 2**, the CDR-H2 comprises the sequence of **SEQ ID NO: 45**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 88**;

3) the CDR-H1 comprises the sequence of **SEQ ID NO: 3**, the CDR-H2 comprises the sequence of **SEQ ID NO: 46**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 89**;

4) the CDR-H1 comprises the sequence of **SEQ ID NO: 4**, the CDR-H2 comprises the sequence of **SEQ ID NO: 47**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 90**;

5) the CDR-H1 comprises the sequence of **SEQ ID NO: 5**, the CDR-H2 comprises the sequence of **SEQ ID NO: 48**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 91**;

6) the CDR-H1 comprises the sequence of **SEQ ID NO: 6**, the CDR-H2 comprises the sequence of **SEQ ID NO: 49**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 92**;

7) the CDR-H1 comprises the sequence of **SEQ ID NO: 7**, the CDR-H2 comprises the sequence of **SEQ ID NO: 50**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 93**;

8) the CDR-H1 comprises the sequence of **SEQ ID NO: 8**, the CDR-H2 comprises the sequence of **SEQ ID NO: 51**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 94**;

9) the CDR-H1 comprises the sequence of **SEQ ID NO: 9**, the CDR-H2 comprises the sequence of **SEQ ID NO: 52**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 95**;

10) the CDR-H1 comprises the sequence of **SEQ ID NO: 10**, the CDR-H2 comprises the sequence of **SEQ ID NO: 53**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 96**;

11) the CDR-H1 comprises the sequence of **SEQ ID NO: 11**, the CDR-H2 comprises the sequence of **SEQ ID NO: 54**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 97**;

12) the CDR-H1 comprises the sequence of **SEQ ID NO: 12**, the CDR-H2 comprises the sequence of **SEQ ID NO: 55**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 98**;

13) the CDR-H1 comprises the sequence of **SEQ ID NO: 13**, the CDR-H2 comprises the sequence of **SEQ ID NO: 56**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 99**;

14) the CDR-H1 comprises the sequence of **SEQ ID NO: 14**, the CDR-H2 comprises the sequence of **SEQ ID NO: 57**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 100**;

15) the CDR-H1 comprises the sequence of **SEQ ID NO: 15**, the CDR-H2 comprises the sequence of **SEQ ID NO: 58**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 101**;

16) the CDR-H1 comprises the sequence of **SEQ ID NO: 16**, the CDR-H2 comprises the sequence of **SEQ ID NO: 59**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 102**;

17) the CDR-H1 comprises the sequence of **SEQ ID NO: 17**, the CDR-H2 comprises the sequence of **SEQ ID NO: 60**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 103**;

18) the CDR-H1 comprises the sequence of **SEQ ID NO: 18**, the CDR-H2 comprises the sequence of **SEQ ID NO: 61**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 104**;

19) the CDR-H1 comprises the sequence of **SEQ ID NO: 19**, the CDR-H2 comprises the sequence of **SEQ ID NO: 62**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 105**;

20) the CDR-H1 comprises the sequence of **SEQ ID NO: 20**, the CDR-H2 comprises the sequence of **SEQ ID NO: 63**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 106**;

21) the CDR-H1 comprises the sequence of **SEQ ID NO: 21**, the CDR-H2 comprises the sequence of **SEQ ID NO: 64**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 107**;

22) the CDR-H1 comprises the sequence of **SEQ ID NO: 22**, the CDR-H2 comprises the sequence of **SEQ ID NO: 65**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 108**;

23) the CDR-H1 comprises the sequence of **SEQ ID NO: 23**, the CDR-H2 comprises the sequence of **SEQ ID NO: 66**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 109**;

24) the CDR-H1 comprises the sequence of **SEQ ID NO: 24**, the CDR-H2 comprises the sequence of **SEQ ID NO: 67**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 110**;

25) the CDR-H1 comprises the sequence of **SEQ ID NO: 25**, the CDR-H2 comprises the sequence of **SEQ ID NO: 68**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 111;**

26) the CDR-H1 comprises the sequence of **SEQ ID NO: 26**, the CDR-H2 comprises the sequence of **SEQ ID NO: 69**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 112**;

27) the CDR-H1 comprises the sequence of **SEQ ID NO: 27**, the CDR-H2 comprises the sequence of **SEQ ID NO: 70**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 113**;

28) the CDR-H1 comprises the sequence of **SEQ ID NO: 28**, the CDR-H2 comprises the sequence of **SEQ ID NO: 71**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 114**;

29) the CDR-H1 comprises the sequence of **SEQ ID NO: 29**, the CDR-H2 comprises the sequence of **SEQ ID NO: 72**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 115**;

30) the CDR-H1 comprises the sequence of **SEQ ID NO: 30**, the CDR-H2 comprises the sequence of **SEQ ID NO: 73**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 116**;

31) the CDR-H1 comprises the sequence of **SEQ ID NO: 31**, the CDR-H2 comprises the sequence of **SEQ ID NO: 74**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 117**;

32) the CDR-H1 comprises the sequence of **SEQ ID NO: 32**, the CDR-H2 comprises the sequence of **SEQ ID NO: 75**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 118**;

33) the CDR-H1 comprises the sequence of **SEQ ID NO: 33**, the CDR-H2 comprises the sequence of **SEQ ID NO: 76**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 119**;

34) the CDR-H1 comprises the sequence of **SEQ ID NO: 34**, the CDR-H2 comprises the sequence of **SEQ ID NO: 77**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 120**;

35) the CDR-H1 comprises the sequence of **SEQ ID NO: 35**, the CDR-H2 comprises the sequence of **SEQ ID NO: 78**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 121**;

36) the CDR-H1 comprises the sequence of **SEQ ID NO: 36**, the CDR-H2 comprises the sequence of **SEQ ID NO: 79**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 122**;

37) the CDR-H1 comprises the sequence of **SEQ ID NO: 37**, the CDR-H2 comprises the sequence of **SEQ ID NO: 80**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 123**;

38) the CDR-H1 comprises the sequence of **SEQ ID NO: 38**, the CDR-H2 comprises the sequence of **SEQ ID NO: 81**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 124**;

39) the CDR-H1 comprises the sequence of **SEQ ID NO: 39**, the CDR-H2 comprises the sequence of **SEQ ID NO: 82**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 125**;

40) the CDR-H1 comprises the sequence of **SEQ ID NO: 40**, the CDR-H2 comprises the sequence of **SEQ ID NO: 83**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 126**;

41) the CDR-H1 comprises the sequence of **SEQ ID NO: 41**, the CDR-H2 comprises the sequence of **SEQ ID NO: 84**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 127**;

42) the CDR-H1 comprises the sequence of **SEQ ID NO: 42**, the CDR-H2 comprises the sequence of **SEQ ID NO: 85**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 128**; or

43) the CDR-H1 comprises the sequence of **SEQ ID NO: 43**, the CDR-H2 comprises the sequence of **SEQ ID NO: 86**, and the CDR-H3 comprises the sequence of **SEQ ID NO: 129**.

3. The CEA6 binding polypeptide of claim 1 or 2, wherein the heavy chain variable domain comprises an amino acid sequence having at least 90%, 95%, 99%, or 100% sequence identity to any sequence selected from **SEQ ID NOs: 130-172**.

4. The CEA6 binding polypeptide of any one of claims 1-3, wherein the CEA6 binding polypeptide is humanized.

5. The CEA6 binding polypeptide of any one of claims 1-4, wherein the CEA6 binding polypeptide is a single domain antibody (sdAb).

6. A chimeric antigen receptor (CAR) comprising the CEA6 binding polypeptide of any one of claims 1-5.

7. A chimeric antigen receptor (CAR) cell comprising the CAR of claim 6.

8. The CAR cell of claim 7, wherein the CAR cell is a CAR T cell.

9. The CAR cell of claim 7 or 8, wherein the CAR cell comprises at least two binding polypeptides and the CAR cell is a multivalent CAR cell.

10. The CAR cell of any one of claims 7-9, wherein the CAR cell is derived from a subject or from a cell line.

11. The CAR cell of claim 10, wherein the subject has a cancer.

12. The CAR cell of claim 11, wherein the cancer is breast cancer, colorectal cancer, kidney cancer, liver cancer, lung cancer, brain cancer, pancreatic cancer, bladder cancer, testicular cancer, prostate cancer, gastric cancer, ovarian cancer, head and neck cancer, gallbladder cancer, a hematologic malignancy, or any combination thereof.

13. A nucleic acid that encodes for a polypeptide comprising a sequence having at least 90%, 95%, 99%, or 100% sequence identity to the CEA6 binding polypeptide of any one of claims 1-5 or the CAR of claim 6.

14. A method of treating a cancer in a subject in need thereof, comprising administering the CAR cell of any one of claims 7-12.

15. The method of claim 14, wherein the chimeric antigen receptor cell is autologous or allogeneic to the subject.

16. The method of claim 14 or 15, wherein the subject is a mammal.

17. The method of any one of claims 14-16, wherein the subject is a human.

18. The method of any one of claims 14-17, wherein the cancer is breast cancer, colorectal cancer, kidney cancer, liver cancer, lung cancer, brain cancer, pancreatic cancer, bladder cancer, testicular cancer, prostate cancer, gastric cancer, ovarian cancer, head and neck cancer, gallbladder cancer, a hematologic malignancy, or any combination thereof.

19. The method of any one of claims 14-18, wherein the chimeric antigen receptor cell is administered parenterally.

CDR-H1

SEQ6	GYILSFNF	8
SEQ27	GYILSFNF	8
SEQ17	GRIYSFNF	8
SEQ39	GPIYSSNF	8
SEQ7	GSIFSSNF	8
SEQ29	GSIFSSNF	8
SEQ32	GSIFSFNT	8
SEQ38	GSIFSFNT	8
SEQ9	GSISSSNF	8
SEQ15	GRTFSYNF	8
SEQ37	GPTFSYNF	8
SEQ18	GPTFSTNF	8
SEQ22	GSTFGLNF	8
SEQ16	GSIINTDA	8
SEQ11	GRISDIYY	8
SEQ34	GRISDIYY	8
SEQ3	GSISGSYY	8
SEQ28	GSISGSYY	8
SEQ31	GSISGSYY	8
SEQ36	GSISGSYY	8
SEQ43	GSISSSYY	8
SEQ4	GSISSPYV	8
SEQ30	GSISSPYV	8
SEQ8	GSILSPYV	8
SEQ20	GPISGSYF	8
SEQ14	GSIYNSYL	8
SEQ1	GSIISYYR	8
SEQ24	GSIXSPYY	8
SEQ5	GSIXSPDY	8
SEQ2	GFIYSPYF	8
SEQ26	GFIYSPYF	8
SEQ41	GFIYSFYF	8
SEQ42	GFIYSFYF	8
SEQ19	GSIISPYY	8
SEQ33	GSIXSPYY	8
SEQ40	GFIGSDYD	8
SEQ12	GFIFDAYD	8
SEQ35	GFIFDAYD	8
SEQ13	GSISNIYD	8
SEQ23	GSISNIYF	8
SEQ10	GSIFDLYY	8
SEQ21	GSIFDSYF	8
SEQ25	GSIFNSYY	8

CDR-H2

SEQ55	~ITSDGDST~	8
SEQ78	~ITSDGDST~	8
SEQ44	~ISSDNSST~	8
SEQ45	~ISSDNSST~	8
SEQ47	~ISRTGGST~	8
SEQ48	~ISRTGGST~	8
SEQ49	~ISRTGGST~	8
SEQ51	~ISPTGGST~	8
SEQ52	~ISSDNSST~	8
SEQ56	~ISSDNSST~	8
SEQ58	~ISSDNSST~	8
SEQ59	~ISRTGGST~	8
SEQ60	~ISRTGGST~	8
SEQ62	~ISRTGGST~	8
SEQ64	~ISRTGGST~	8
SEQ65	~ISSDNSST~	8
SEQ67	~ISSDNSST~	8
SEQ68	~ISSDNSST~	8
SEQ69	~ISRTGGST~	8
SEQ70	~ISRTGGST~	8
SEQ73	~ISRTGGST~	8
SEQ76	~ISRTGGST~	8
SEQ80	~ISRTGGST~	8
SEQ82	~ISPTGGST~	8
SEQ83	~ISSDNSST~	8
SEQ84	~ISRTGGST~	8
SEQ85	~ISRTGGST~	8
SEQ54	SINSYAGITY	10
SEQ77	SINSYAGITY	10
SEQ53	TINSLDAITY	10
SEQ61	~IUSFGSST~	8
SEQ46	TINSPNDITY	10
SEQ71	TINSPNDITY	10
SEQ74	TINSFGGITY	10
SEQ79	TINSFGGITY	10
SEQ86	TINSFGGITY	10
SEQ57	~ISSFGGIT~	8
SEQ63	~INTIASST~	8
SEQ50	TISWNSDSTY	10
SEQ72	TISWNSDSTY	10
SEQ75	TISWSSDSTY	10
SEQ81	TISWSSDSTY	10
SEQ66	~INSISGST~	8

CDR-H3

SEQ98	-----------AVLPFPWIRQYGY-------------	13
SEQ121	-----------AVLPFPWIRQYGY-------------	13
SEQ95	-------------------------VTQSWASGKIP----SY	13
SEQ92	--A---------SARGPSLTVPDY------------------	13
SEQ113	AA---------SARGPSLTVPDY------------------	14
SEQ109	-------------AVPPWGGYDAYRSDETFIYYSY	22
SEQ100	AA---WPYYQDRATLLAHNY---------------------	17
SEQ101	----------AAIEDVLNHNY--------------------	11
SEQ123	-----------AIEDVLNHNY--------------------	10
SEQ91	--A-------ANTALSILDLTAPDY----------------	16
SEQ90	--A-------ANTAL-----TWYNYDY--------------	13
SEQ116	AA-----ANTAL-----TWYNYDY--------------	14
SEQ89	--AGIRMIT------IEHRNKYSY-----------------	16
SEQ112	AAGIRMIT-------IEHRNKYSY-----------------	17
SEQ127	--AGIRMIT-------IEHRNKYSY----------------	16
SEQ129	--AGIPNIT-------IEHWNKYSY----------------	16
SEQ105	AVTRTAIT-------LEQVGQGSY-----------------	17
SEQ119	AVTRTAIT-------LEQVGQGSY-----------------	17
SEQ103	AAGVMVIV-----VTINQRYDY-------------------	17
SEQ125	--AGVMVIV-----VTINQRYDY------------------	16
SEQ107	AAIYSPPI-------QMSNRSFDY-----------------	17
SEQ87	--AAARSMI-------TSSLAYYDY----------------	16
SEQ99	AASPVPPL-------TLAEQYYDY-----------------	17
SEQ110	AASPA---LVS---VQQHEGFNY------------------	17
SEQ111	AASGA----TVC----VADASSDNY----------------	17
SEQ106	---AAKENDKLATTYSY------------------------	14
SEQ104	--AVVVSPSTEYGSIVIYGYDY-------------------	20
SEQ102	AAGXR----------VSPTLNFDY-----------------	14
SEQ108	AAMER----------SKPSLDSDY-----------------	14
SEQ94	--VAMR---------NHITLDSDY-----------------	13
SEQ126	--AAST---------TFSSHN---Y----------------	11
SEQ93	----ANT---------GRSASSHNY----------------	12
SEQ115	--AANT----------GRSASSHNY----------------	13
SEQ118	--AANT----------GHSASSHNY----------------	13
SEQ124	----ANT----------GHSASSHNY---------------	12
SEQ88	----ANT-----------GHSDPSY----------------	10
SEQ114	--AANT-----------GHSDPSY-----------------	11
SEQ117	--AANT-----------GHSDPSY-----------------	11
SEQ122	----ANT-----------GHSDPNY----------------	10
SEQ128	----ANT-----------GHSDPNY----------------	10
SEQ96	--AANT-----------GYSSLSY-----------------	11
SEQ97	--AANT-----------GPSSLSY-----------------	11
SEQ120	----ANT-----------GSSSLSY----------------	10

FIG. 1

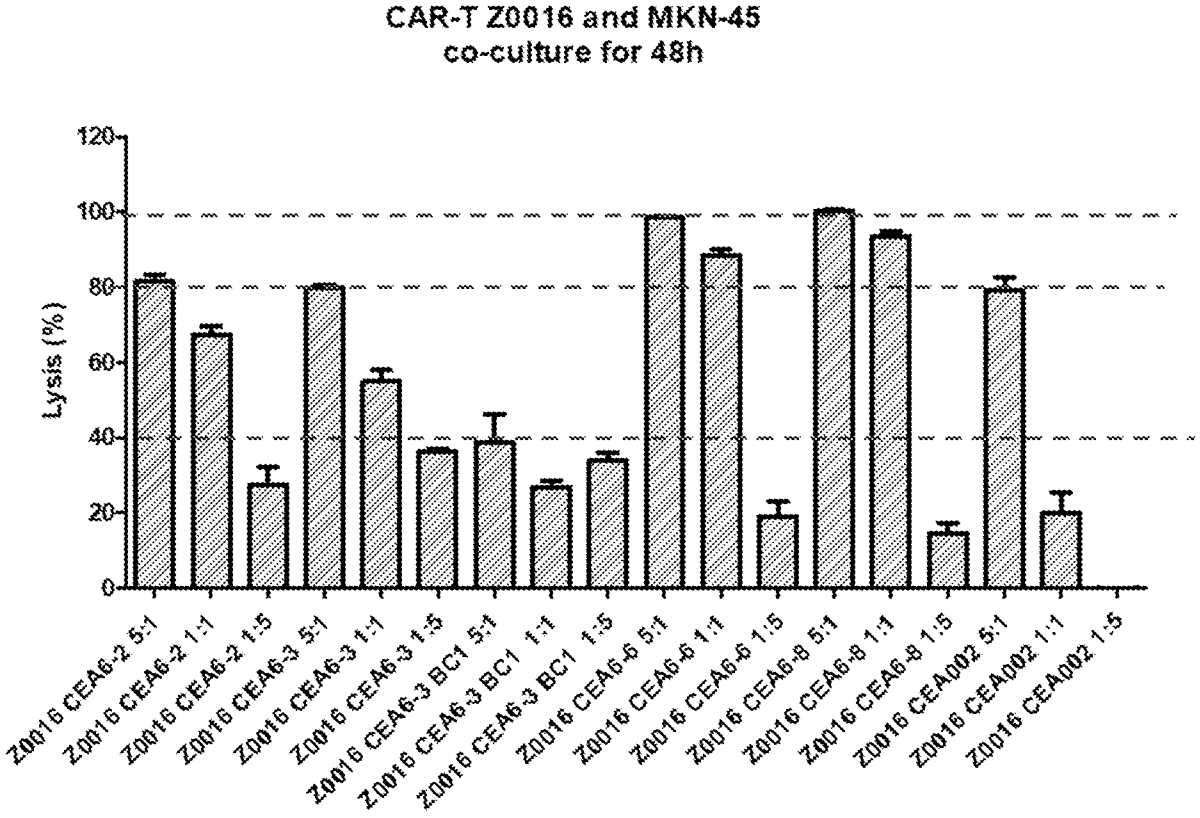

CAR-T Z0016 and MKN-45
co-culture for 48h
Lysis (%)
120
100
80
60
40
20
0
Z0016 CEA6-2 5:1
Z0016 CEA6-2 1:1
Z0016 CEA6-2 1:5
Z0016 CEA6-3 5:1
Z0016 CEA6-3 1:1
Z0016 CEA6-3 1:5
Z0016 CEA6-3 BC1 5:1
Z0016 CEA6-3 BC1 1:1
Z0016 CEA6-3 BC1 1:5
Z0016 CEA6-6 5:1
Z0016 CEA6-6 1:1
Z0016 CEA6-6 1:5
Z0016 CEA6-8 5:1
Z0016 CEA6-8 1:1
Z0016 CEA6-8 1:5
Z0016 CEA002 5:1
Z0016 CEA002 1:1
Z0016 CEA002 1:5

FIG. 2A

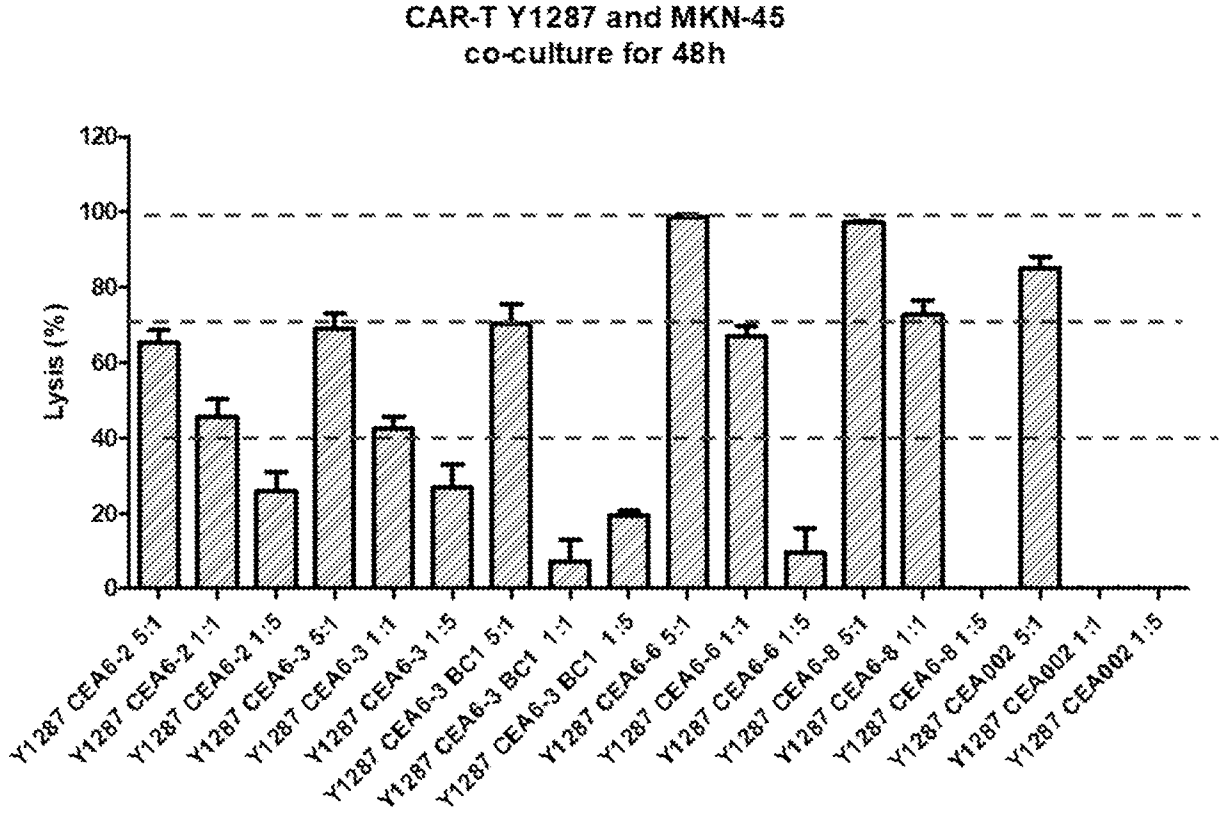

CAR-T Y1287 and MKN-45
co-culture for 48h
Lysis (%)
120
100
80
60
40
20
0
Y1287 CEA6-2 5:1
Y1287 CEA6-2 1:1
Y1287 CEA6-2 1:5
Y1287 CEA6-3 5:1
Y1287 CEA6-3 1:1
Y1287 CEA6-3 1:5
Y1287 CEA6-3 BC1 5:1
Y1287 CEA6-3 BC1 1:1
Y1287 CEA6-3 BC1 1:5
Y1287 CEA6-6 5:1
Y1287 CEA6-6 1:1
Y1287 CEA6-6 1:5
Y1287 CEA6-8 5:1
Y1287 CEA6-8 1:1
Y1287 CEA6-8 1:5
Y1287 CEA002 5:1
Y1287 CEA002 1:1
Y1287 CEA002 1:5

FIG. 2B

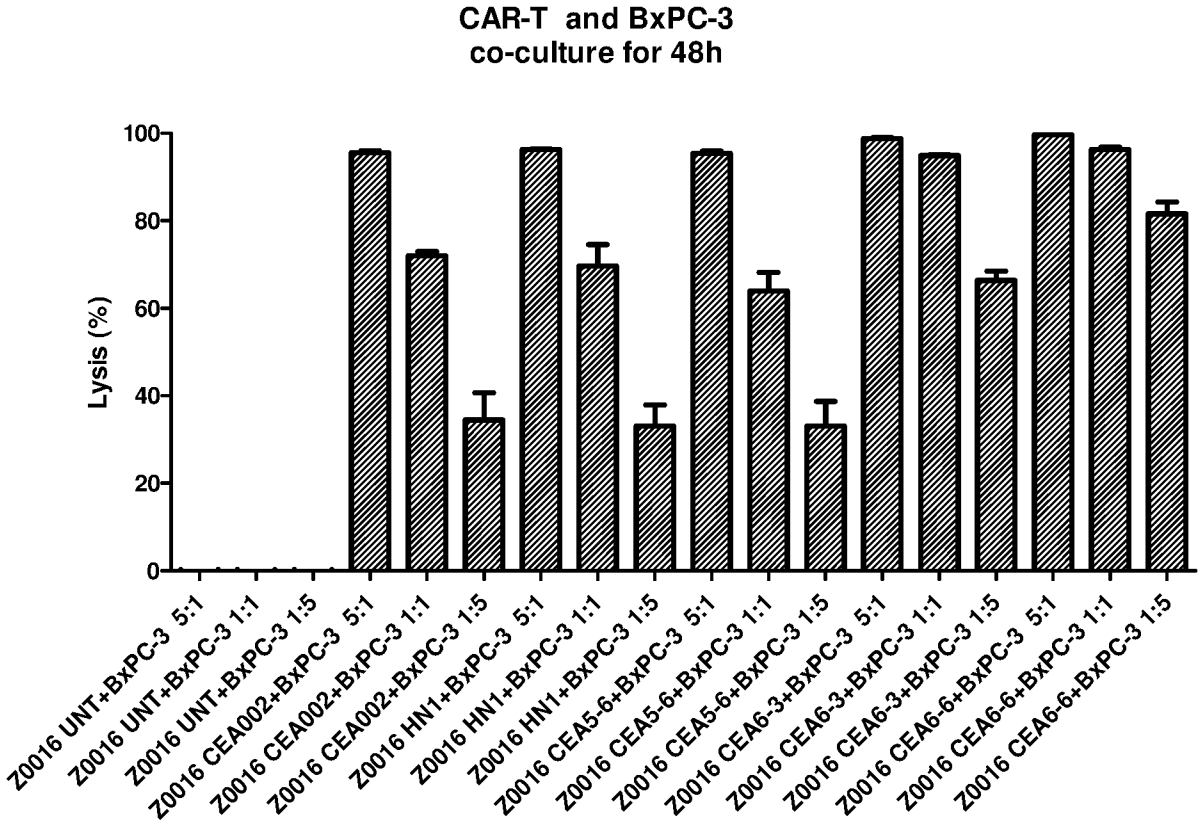

CAR-T and BxPC-3
co-culture for 48h
Lysis (%)
100
80
60
40
20
0
Z0016 UNT+BxPC-3 5:1
Z0016 UNT+BxPC-3 1:1
Z0016 UNT+BxPC-3 1:5
Z0016 CEA002+BxPC-3 5:1
Z0016 CEA002+BxPC-3 1:1
Z0016 CEA002+BxPC-3 1:5
Z0016 HN1+BxPC-3 5:1
Z0016 HN1+BxPC-3 1:1
Z0016 HN1+BxPC-3 1:5
Z0016 CEA5-6+BxPC-3 5:1
Z0016 CEA5-6+BxPC-3 1:1
Z0016 CEA5-6+BxPC-3 1:5
Z0016 CEA6-3+BxPC-3 5:1
Z0016 CEA6-3+BxPC-3 1:1
Z0016 CEA6-3+BxPC-3 1:5
Z0016 CEA6-6+BxPC-3 5:1
Z0016 CEA6-6+BxPC-3 1:1
Z0016 CEA6-6+BxPC-3 1:5

FIG. 2C

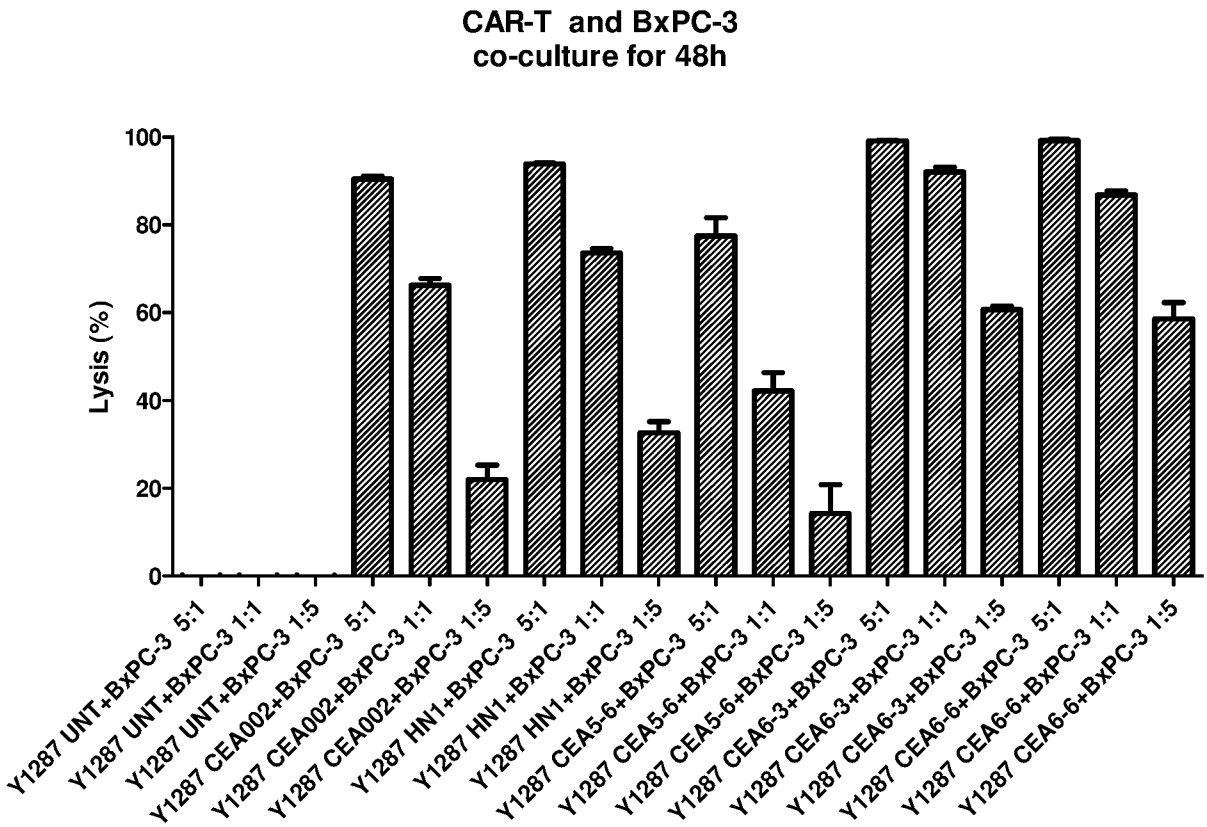

CAR-T and BxPC-3
co-culture for 48h
Lysis (%)
100
80
60
40
20
0
Y1287 UNT+BxPC-3 5:1
Y1287 UNT+BxPC-3 1:1
Y1287 UNT+BxPC-3 1:5
Y1287 CEA002+BxPC-3 5:1
Y1287 CEA002+BxPC-3 1:1
Y1287 CEA002+BxPC-3 1:5
Y1287 HN1+BxPC-3 5:1
Y1287 HN1+BxPC-3 1:1
Y1287 HN1+BxPC-3 1:5
Y1287 CEA5-6+BxPC-3 5:1
Y1287 CEA5-6+BxPC-3 1:1
Y1287 CEA5-6+BxPC-3 1:5
Y1287 CEA6-3+BxPC-3 5:1
Y1287 CEA6-3+BxPC-3 1:1
Y1287 CEA6-3+BxPC-3 1:5
Y1287 CEA6-6+BxPC-3 5:1
Y1287 CEA6-6+BxPC-3 1:1
Y1287 CEA6-6+BxPC-3 1:5

FIG. 2D

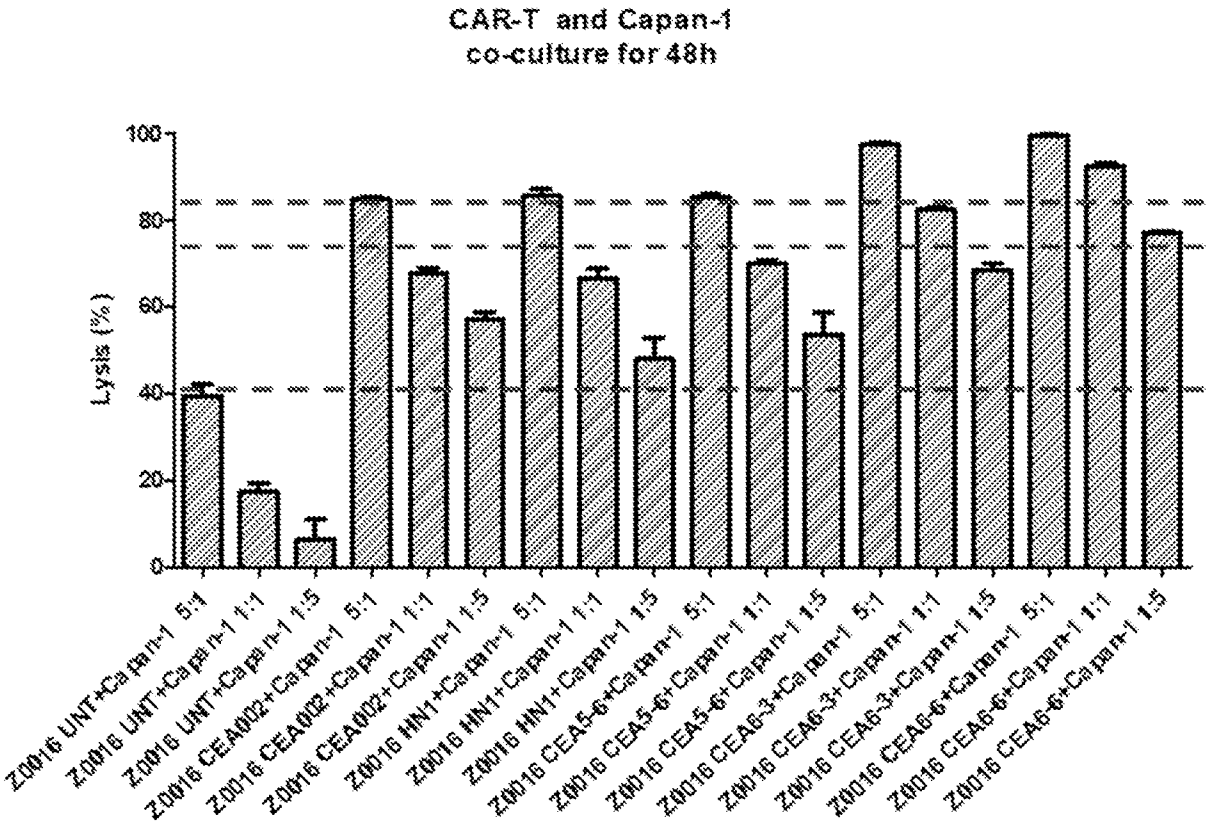

CAR-T and Capan-1
co-culture for 48h
Lysis (%)
100
80
60
40
20
0
Z0016 UNT+Capan-1 5:1
Z0016 UNT+Capan-1 1:1
Z0016 UNT+Capan-1 1:5
Z0016 CEA6D2+Capan-1 5:1
Z0016 CEA6D2+Capan-1 1:1
Z0016 CEA6D2+Capan-1 1:5
Z0016 HN1+Capan-1 5:1
Z0016 HN1+Capan-1 1:1
Z0016 HN1+Capan-1 1:5
Z0016 CEA5-6+Capan-1 5:1
Z0016 CEA5-6+Capan-1 1:1
Z0016 CEA5-6+Capan-1 1:5
Z0016 CEA6-3+Capan-1 5:1
Z0016 CEA6-3+Capan-1 1:1
Z0016 CEA6-3+Capan-1 1:5
Z0016 CEA6-6+Capan-1 5:1
Z0016 CEA6-6+Capan-1 1:1
Z0016 CEA6-6+Capan-1 1:5

FIG. 2E

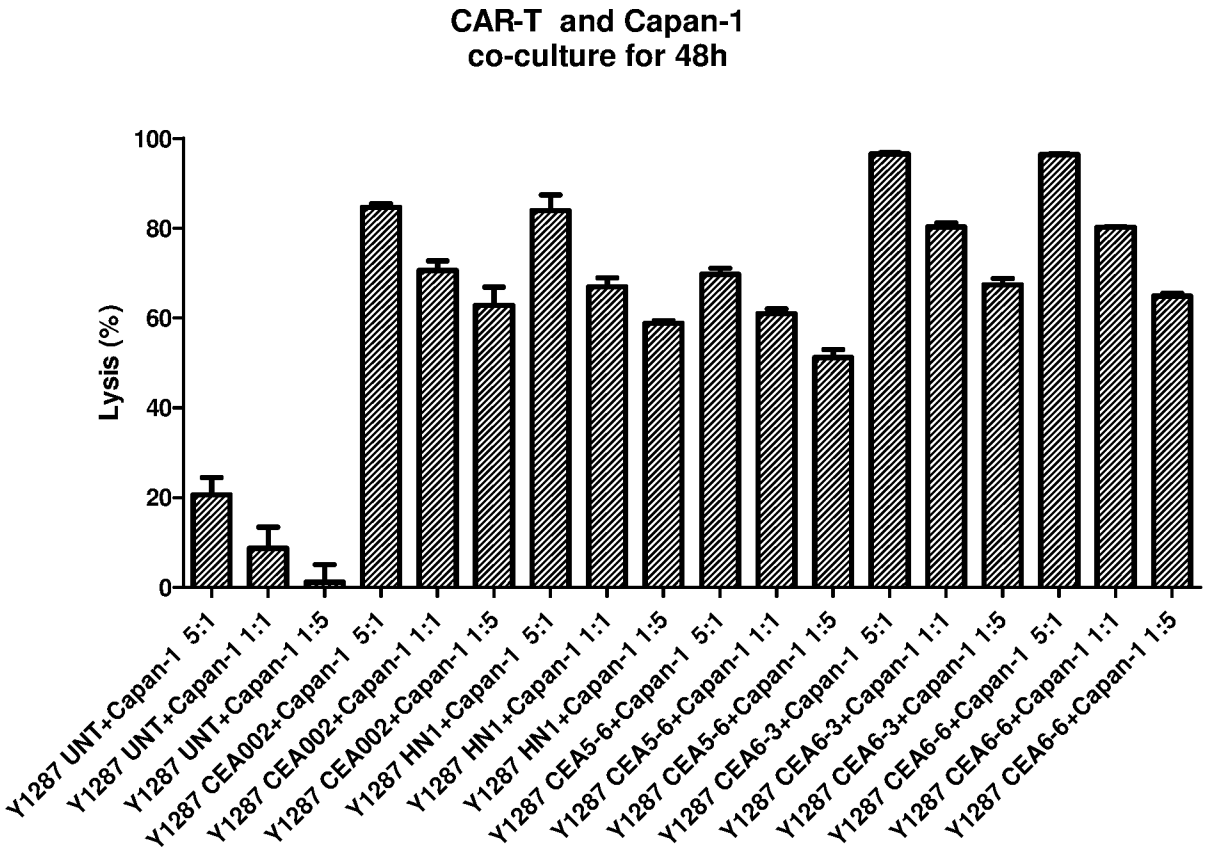
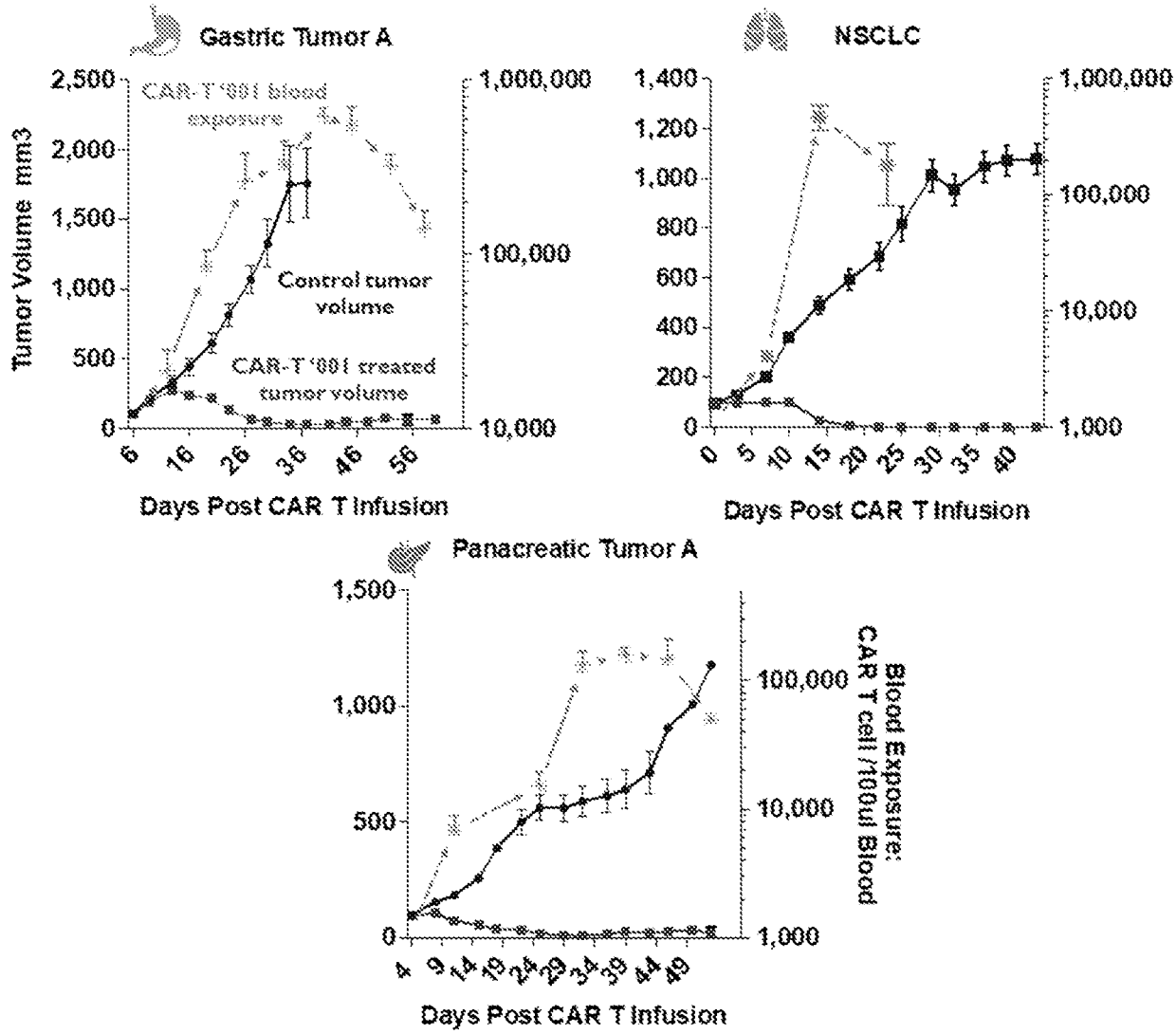

CAR-T and Capan-1
co-culture for 48h
Lysis (%)
100
80
60
40
20
0
Y1287 UNT+Capan-1 5:1
Y1287 UNT+Capan-1 1:1
Y1287 UNT+Capan-1 1:5
Y1287 CEA002+Capan-1 5:1
Y1287 CEA002+Capan-1 1:1
Y1287 CEA002+Capan-1 1:5
Y1287 HN1+Capan-1 5:1
Y1287 HN1+Capan-1 1:1
Y1287 HN1+Capan-1 1:5
Y1287 CEA5-6+Capan-1 5:1
Y1287 CEA5-6+Capan-1 1:1
Y1287 CEA5-6+Capan-1 1:5
Y1287 CEA6-3+Capan-1 5:1
Y1287 CEA6-3+Capan-1 1:1
Y1287 CEA6-3+Capan-1 1:5
Y1287 CEA6-6+Capan-1 5:1
Y1287 CEA6-6+Capan-1 1:1
Y1287 CEA6-6+Capan-1 1:5

FIG. 2F



Note: each tumor type study included NSG mice (n=5)

FIG. 3A

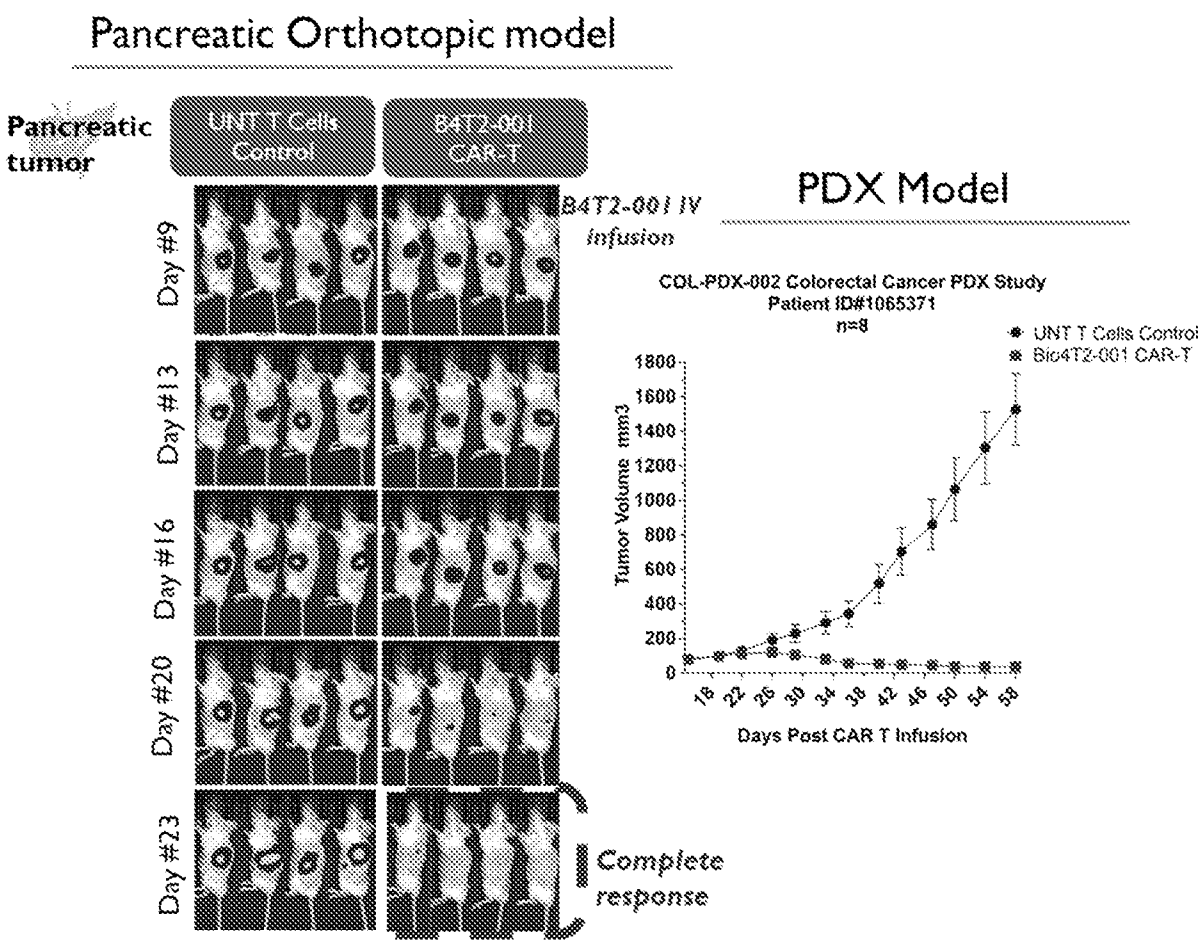

Pancreatic Orthotopic model
Pancreatic tumor
UNT T Cells Control
B4T2-001 CAR-T
Day #9
Day #13
Day #16
Day #20
Day #23
B4T2-001 IV Infusion
Complete response
PDX Model
COL-PDX-002 Colorectal Cancer PDX Study
Patient ID#1065371
n=8
UNT T Cells Control
Bio4T2-001 CAR-T
Tumor Volume mm3
1800
1600
1400
1200
1000
800
600
400
200
0
16
22
26
30
34
38
42
46
50
54
58
Days Post CAR T Infusion

FIG. 3B

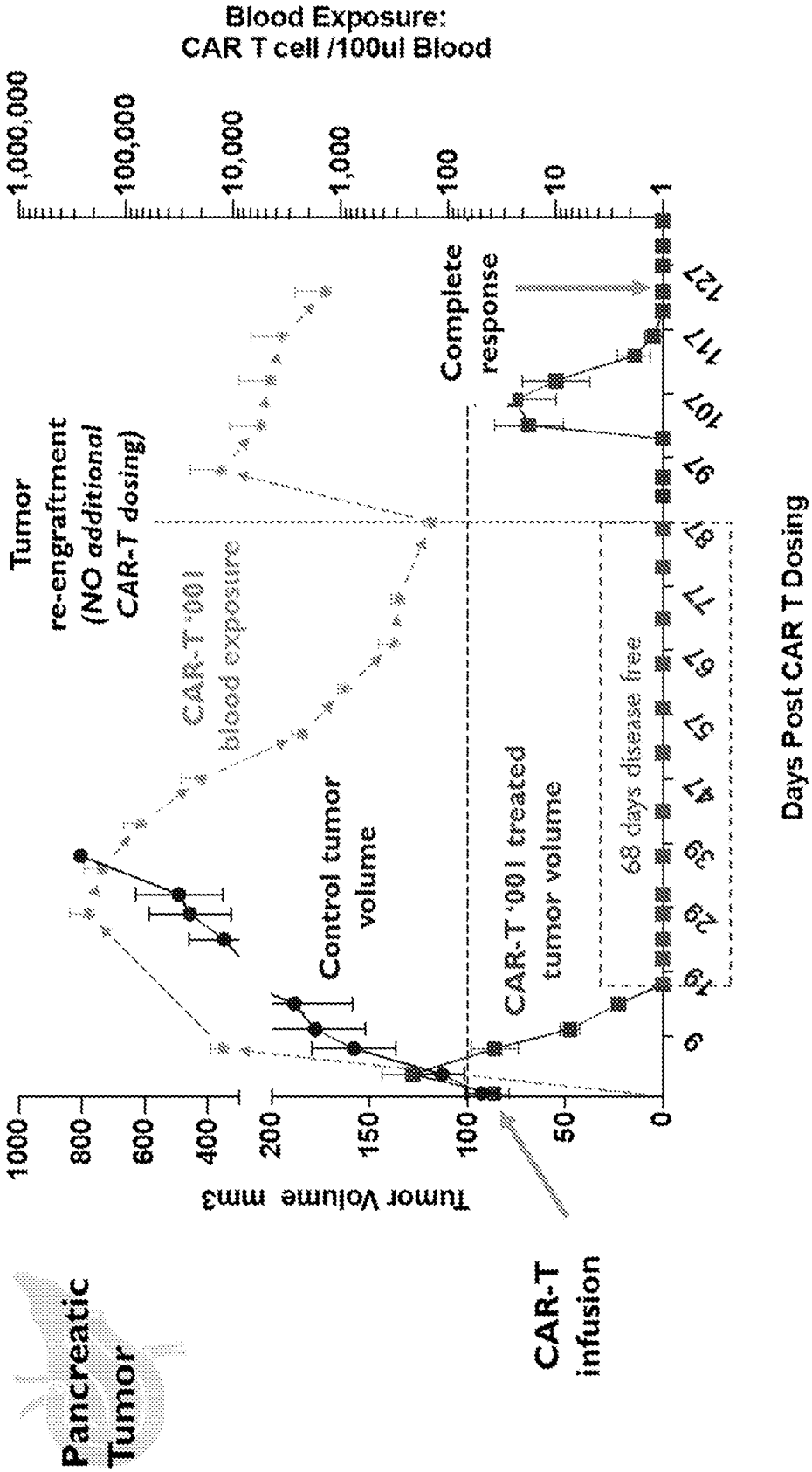

Blood Exposure:
CAR T cell /100ul Blood
1,000,000
100,000
10,000
1,000
100
10
1
Tumor
re-engraftment
(NO additional
CAR-T dosing)
CAR-T '001
blood exposure
Complete
response
Control tumor
volume
CAR-T '001 treated
tumor volume
68 days disease free
Tumor Volume mm3
1000
800
600
400
200
150
100
50
0
CAR-T
infusion
Pancreatic
Tumor
6
27
34
41
48
55
62
69
76
83
90
97
104
111
118
125
Days Post CAR T Dosing

FIG. 4

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 21/63812

A. CLASSIFICATION OF SUBJECT MATTER
IPC - C07K 16/28, A61P 35/00 (2022.01)
CPC - A61K 39/3955, A61P 35/00, C07K 16/2803, C12N 15/63, C07K 2317/509

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)
See Search History document

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

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2020/0181261 A1 (BAYER PHARMA AKTIENGESELLSCHAFT) 11 June 2020 (11.06.2020) Claim 1, para [0081]	1-3
A	US 2011/0177095 A1 (HARDING et al.) 21 July 2011 (21.07.2011) abstract, SEQ ID NO: 439	1-3
A	WO 2019/178316 A1 (GENENTECH, INC.) 19 September 2019 (19.09.2019) abstract, SEQ ID NO: 73	1-3
A	US 2010/0022452 A1 (SILENCE) 28 January 2010 (28.01.2010) abstract, para [0054], [0076], SEQ ID NOs: 31, 213	1-3
A	WO 2019/126399 A1 (SURROZEN INC.) 27 June 2019 (27.06.2019) abstract, Table 1A, SEQ ID NO: 392	1-3
A	US 2005/0053608 A1 (WEBER et al.) 10 March 2005 (10.03.2005) abstract, para [0345], SEQ ID NO: 96	1-3

☐ 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
"D" document cited by the applicant in the international application
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
25 April 2022	MAY 12 2022
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-8300	Authorized officer Kari Rodriquez Telephone No. PCT Helpdesk: 571-272-4300

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 21/63812

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

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

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

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

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

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

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

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

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

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

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 21/63812

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

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

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

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

3. ☒ Claims Nos.: 4-19
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

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

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

- see extra sheet for Box No. III Observations where unity of invention is lacking -

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

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

3. ☒ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
1-3 limited to SEQ ID NOs: 1, 17, 44, 60, 87, 103, 130, 146

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

Remark on Protest

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

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

☐ No protest accompanied the payment of additional search fees.

Continuation of:
Box No. III. Observations where unity of invention is lacking

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Groups I+: Claims 1-3, drawn to a carcinoembryonic antigen 6 (CEA6) binding polypeptide comprising an immunoglobulin heavy chain variable domain comprising a CDR-H1, CDR-H2, and CDR-H3. The composition will be searched to the extent that the CDR-H1, CDR-H2, and CDR-H3 encompasses SEQ ID NOs: 1, 44, and 87, respectively; and VHH of SEQ ID NO: 130. It is believed that claims 1-3 encompass this first named invention, and thus these claims will be searched without fee to the extent that they encompass SEQ ID NOs: 1, 44, 87 and 130. Additional CEA6 binding polypeptide(s) will be searched upon the payment of additional fees. Applicants must specify the claims that encompass any additionally elected CEA6 binding polypeptide(s). Applicants must further indicate, if applicable, the claims which encompass the first named invention, if different than what was indicated above for this group. Failure to clearly identify how any paid additional invention fees are to be applied to the "+" group(s) will result in only the first claimed invention to be searched. An exemplary election would be a CEA6 binding polypeptide(s) encompassing CDR-H1, CDR-H2, and CDR-H3 of SEQ ID NOs: 2, 45, and 88, respectively; and VHH of SEQ ID NO: 131 (Claims 1-3).

The inventions listed as Groups I+ do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

No technical features are shared between the amino acid sequences of Groups I+ and, accordingly, these groups lack unity a priori.

Additionally, even if Groups I+ were considered to share the technical features of including: a CEA6 binding polypeptide, these shared technical features are previously disclosed by US 2020/0181261 A1 to Bayer Pharma Aktiengesellschaft (hereinafter "Bayer").

Bayer teaches a carcinoembryonic antigen 6 (CEA6) binding polypeptide comprising an immunoglobulin heavy chain variable domain comprising a CDR-H1, CDR-H2, and CDR-H3 (Claim 1, An isolated anti-CEACAM6 antibody or antigen-binding fragment thereof comprising: xiii. a heavy chain antigen-binding region that comprises an H-CDR1 comprising SEQ ID NO:106, an H-CDR2 comprising SEQ ID NO:107, and an H-CDR3 comprising SEQ ID NO:108.; para [0081], "Functional fragments", "antigen-binding antibody fragments", or "antibody fragments" of the invention include but are not limited to Fab, Fab', Fab'-SH, F(ab')2, and Fv fragments; diabodies; single domain antibodies (Dabs).).

As said technical features were known in the art at the time of the invention, these cannot be considered special technical features that would otherwise unify the groups.

Groups I+ therefore lack unity under PCT Rule 13 because they do not share a same or corresponding special technical feature.

Item 4 (continued):
Claims 4-19 are improper multiple dependent claims because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).